ESSAYS ON MACROECONOMICS AND FORECASTING

by

Natsuki Arai

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

October, 2014

Abstract

This dissertation consists of three essays on macroeconomics and forecasting. Central banks have recently engaged in unconventional monetary policies at the zero lower bound, where their forecasts play a crucial role in signaling the future path of monetary policy. The work presented herein investigates the effect of monetary policy announcements at the zero lower bound in Japan, and evaluates the efficiency of forecasts that are explicitly tied to monetary policy decisions in the United States.

The first chapter investigates the effects of monetary policy announcements at the zero lower bound using Japanese data from 1998 to 2013. I find that the effect of expansionary monetary policy shocks is directly passed on to corporate bond yields, notably for high-grade corporate bond yields. However, the magnitude of estimated pass-through to stock prices and the exchange rate is substantially smaller than in the U.S., and not statistically significant in most cases. Such differences may reflect a higher degree of market segmentation or smaller scope to achieve further accommodation in Japan.

The second chapter evaluates the efficiency of the FOMC's new economic projections. Since 2007, FOMC policymakers have been publishing detailed numerical projections of macroeconomic series over the next three years. By testing whether the revisions to these projections are unpredictable, I find that FOMC's efficiency is generally accepted for inflation, but often rejected for real economic variables, notably for the unemployment rate. The rejection is due to the strong autocorrelation of revisions, which may reflect information rigidity of

FOMC's unemployment projections. The joint efficiency of the entire projection is accepted in most cases.

The third chapter evaluates the efficiency of Fed's Greenbook forecast and uses this evaluation to improve the accuracy of the Greenbook forecast. Recently, Patton and Timmermann (2012) proposed a more powerful kind of forecast efficiency regression at multiple horizons, and showed that it provides evidence against the efficiency of the Fed's Greenbook forecasts. I use their forecast efficiency evaluation to propose a method for adjusting the Greenbook forecasts. Using this method in a real-time out-of-sample forecasting exercise, I find that it gives modest improvements in the accuracy of the forecasts for GDP deflator and CPI, but not for other variables. The improvements are statistically significant in some cases, the magnitude of which can be as high as an 18 percent reduction in the root mean square prediction error.

Keywords:	Unconventional Monetary Policy, Zero Lower Bound, Identification through Heteroscedasticity, Forecast Efficiency, Forecast Revisions, Real-Time Data
JEL classification:	C32, C53, E43, E44, E52, E58
Primary Reader:	Jonathan H. Wright
Secondary Reader:	Jon W. Faust

Acknowledgement

This dissertation could not have been completed without help from countless people through my Ph.D. journey. I want to express my deep appreciation to Jonathan Wright, for his unconditional support, guidance, and broad vision throughout my dissertation. I am very lucky to have him as my primary advisor and a role model in my graduate career. I would also like to thank Jon Faust for his open discussion and brilliant ideas that enlighten my research. I have learned a lot from both advisors and I am truly grateful for them. I would like to thank my dissertation committee member Larry Ball for his helpful advice.

I have benefited from the discussion with David Aadland, Bob Barbera, Chris Carroll, Varanya Chaubey, Ryo Ishida, Takashi Kano, Katsuhiro Oshima, Matt Raskin, Yuya Sasaki, Richard Spady, David Vera, Egon Zakrajsek, and the seminar participants at Hitotsubashi University, Georgetown University, Johns Hopkins University, National Tsing Hua University, National Chengchi University, the 2011 Missouri Economics conference, the 2012 and 2013 Eastern Economic conference, and the 2013 International Symposium of Forecasting.

Graduate school is a long journey and I am grateful for my fellow travelers: Blair Chapman, Yoichi Goto, Burçin Kisacikoglu, Sangah Lee, Chris Martin, Michele Mazzoleni, Rodrigo Sekkel, Ruli Xiao, and Jiae Yoo. Special thanks to Rodrigo Sekkel for his advice both in research and the life in the graduate school.

I would like to thank my parents, Koichiro Arai and Yuko Arai, for their enormous patience and tolerance, and constant encouragement and support. Without their love, I would not complete such a long journey. This dissertation

is dedicated to them.

Last but not least, I would like to thank my wife, An-Chi Wei, for brightening my life by her unconditional love and support. You make my life more joyful.

Table of Contents

List of Tables

List of Figures

Chapter 1

The Effects of Monetary Policy Announcements at the Zero Lower Bound

1.1 Introduction

The effect of unconventional monetary policy at the zero lower bound (ZLB), which includes forward guidance and asset purchases, has been a centerpiece of the debate in macro-finance since many advanced economies reached the ZLB after the financial crisis of 2008. Since the crisis, a number of important contributions have been made regarding this topic.[1] However, the analysis in the literature primarily focuses on the U.S. economy after the crisis, which is restricted by a short sample, and researchers are not sure about the effect of unconventional monetary policy in a different environment.

In this paper, I study the effects of unconventional monetary policy in Japan, which has experienced a substantially longer period at the ZLB, from 1995

[1] For example, see D'Amico and King (2013), Gagnon, Raskin, Remache, and Sack (2011), Krishnamurthy and Vissing-Jorgensen (2011), and Wright (2012) for the Fed's Large-Scale Asset Purchases (LSAPs) programs, and see Joyce, Miles, Scott, and Vayanos (2012) for the Bank of England's asset purchases. For more comprehensive review, see Bernanke (2012c). For the comparison across advanced economies, see Rogers, Scotti, and Wright (2014).

to the present. Specifically, I use the method of identification through heteroscedasticity, which was originally proposed by Rigobon (2003) and Rigobon and Sack (2003, 2004) and has been widely used in the recent literature,[2] to estimate the pass-through of monetary policy shock. Identification is based on the assumption that the variance of monetary policy shocks is particularly high on important announcement days, whereas nothing unusual happens to other shocks on these days.

To assess the stimulative effect of monetary policy on aggregate demand, I focus on the pass-through of monetary policy shock to three financial assets, which are commonly targeted by central banks: corporate bonds, stocks and the exchange rate. First, I study the pass-through to corporate bond yields, because the reduction in the borrowing cost of firms is a key channel through which monetary policy could stimulate aggregate demand. Second, I analyze the pass-through to stock prices, which is relevant because the response of stock prices could increase consumption through the wealth effect. Finally, I evaluate the pass-through to the exchange rate, through which aggregate demand can be boosted via the trade balance.

The results show that there is a stark contrast between the pass-through to corporate bond yields and the pass-through to stock prices and the exchange rate. For corporate bond yields, there is a statistically significant and about one-to-one pass-through, notably for high-grade corporate bond yields. This result is broadly similar to the U.S. estimates in Raskin (2013a). On the other hand, the pass-through to stock prices and the exchange rate is not statistically significant in most cases. It contrasts with the results using the U.S. data

[2]For example, see Gilchrist and Zakrajsek (2013) and Raskin (2013a).

in Kiley (2013), Neely (2013), and Glick and Leduc (2013), which show the statistically significant pass-through to these assets.

To interpret the smaller pass-through to stock prices and the exchange rate in Japan, I discuss two possible explanations. The first explanation is the higher degree of market segmentation in Japan. In other words, the Japanese financial market may have some institutional features that lead to inefficient monetary transmission mechanisms. The second explanation is the smaller scope for the forward guidance policy to provide effective accommodation. More specifically, in the economy that has experienced a prolonged period at the ZLB, expectations about short-term rates are stuck at zero and the intended effect of forward guidance policy could be substantially limited.

In addition, I use a simple event study to analyze the effects of announcements in 2013, to show that these announcements have substantial effect even on stock prices. The announcements in 2013 are associated with the regime change of Bank of Japan's (BOJ) monetary policy to commit to the 2-percent inflation target by 2015. Unlike the previous announcements, these announcements had substantial effects, not only on corporate bonds, but also on stock prices. This difference may be due to the different nature of the BOJ's commitment after 2013; the commitment is open-ended and the BOJ announces that it will do whatever it takes to achieve the target inflation.

Based on the results in this paper and survey measures, I provide a conjecture about the decomposition of the effect of unconventional monetary policy. Given that survey measures of inflation expectations and short-term interest rates have been stable, it is likely that unconventional monetary policy affects the term premium. Furthermore, it is even more likely that unconventional monetary policy affects the term premium in the recent regime change; Inflation expectation increases responding to the announcements while the long-term yields decline, suggesting that we have reduction in the term premium.

Lastly, I provide several robustness checks, to which the main results are generally robust. First, I analyze the pass-through to other financial assets: (1) real estate investment trusts (REIT), (2) credit default swaps (CDS), and (3) the exchange rate of the OECD and Asian economies. Second, I consider subsamples focusing on different programs: 2001-2006, 2006-2010, and 2010-2013. Third, I provide the analysis based on the principal component of government bond yields with different maturities. Last, I use alternative sets of non-announcement days.

The remainder of the paper is organized as follows: Section 1.2 describes the methodology used in the paper, Section 1.3 explains the data and background of Japanese monetary policy, Section 1.4 presents the results and discussion. Section 1.5 concludes.

1.2 Method

This section describes an analytical framework to estimate the pass-through of monetary policy shocks on various financial assets. Based on the standard setup of two simultaneous equations, I present a simple event study and the framework of identification through heteroscedasticity. In addition, I introduce the weak-identification robust confidence set to address the issue of weak identification.

1.2.1 Setup

Consider the system of two simultaneous equations between the change in the interest rate and the growth rate of the asset price, Δi_t and Δs_t. The notation follows Rigobon and Sack (2004):

$$\Delta i_t = \beta \Delta s_t + \gamma X_t + \varepsilon_t, \tag{1.1}$$

$$\Delta s_t = \alpha \Delta i_t + \delta X_t + \eta_t, \tag{1.2}$$

where X_t is a common exogenous shock that simultaneously affects both the interest rate and the asset price, ε_t is a monetary policy shock, and η_t is a shock to the asset price.

In this system, I primarily focus on estimating the parameter α because it indicates how much monetary policy shocks affect asset prices through the changes in the interest rate. However, the OLS estimate of the pass-through, α, is biased since both variables, Δi_t and Δs_t, are simultaneously determined in the system.[3]

1.2.2 Event Study

An event study is a simple way to estimate the pass-through using a directly measured monetary policy surprise. By picking the important announcements and regarding them as complete surprises, we can use the corresponding changes in the asset prices to estimate the effect of monetary policy shocks. Gagnon, Raskin, Remache, and Sack (2011) directly measured the effect of the Fed's Large-Scale Asset Purchases (LSAP) by assuming that the announcements about

[3] For the derivation of the OLS estimate and its bias, see Appendix A.1.

the LSAP were complete surprises, and added up the changes on the announcement days. Though the event study is based on the strong assumption that no other material news came within the announcement window, it can provide useful benchmark results.

1.2.3 Identification through Heteroscedasticity

To obtain a consistent estimate of the pass-through under weaker assumptions, we employ a scheme called identification through heteroscedasticity, proposed by Rigobon (2003) and Rigobon and Sack (2003, 2004). Essentially, it uses the shift of the variances of endogenous variables between the announcement days and non-announcement days as instruments for the identification.

I introduce some notation and assumptions to describe this scheme of identification. First, I denote a subset of the policy announcement days as A and a subset of the non-announcement days as $\bar{A}$.[4] Second, I denote the number of announcement days and non-announcement days as T_A and $T_{\bar{A}}$, and thus the total number of days as $T \equiv T_A + T_{\bar{A}}$. Finally, I assume that the variance of monetary policy shock is larger on the announcement days than on the non-announcement days, but the variance of other shocks are the same across these two sets of days. Under this assumption, the difference of the conditional variance-covariance matrices in these two sets of days, $\mathbf{\Omega}_A$ and $\mathbf{\Omega}_{\bar{A}}$, only depends on the variance of monetary policy shocks. Specifically, we can compute

[4] Unlike Rigobon and Sack (2004), all business days that do not belong to A are treated as the non-announcement days. To make this point clear, I denote a subset of the non-announcement days as $\bar{A}$. The results using alternative sets of the non-announcement days are presented in Section 1.4.5.4.

the difference of the variances, $\boldsymbol{\Delta\Omega}$, as follows:

$$\boldsymbol{\Delta\Omega} \equiv \boldsymbol{\Omega}_A - \boldsymbol{\Omega}_{\bar{A}} = \frac{\sigma^2_{\varepsilon|A} - \sigma^2_{\varepsilon|\bar{A}}}{(1-\alpha\beta)^2} \begin{pmatrix} 1 & \alpha \\ \alpha & \alpha^2 \end{pmatrix}, \tag{1.3}$$

where $\sigma^2_{\varepsilon|A}$ and $\sigma^2_{\varepsilon|\bar{A}}$ are the conditional variances of monetary policy shocks on the announcement days and the non-announcement days, respectively.[5] This is because the effects of other shocks are cancelled out by taking the difference between the announcement days and non-announcement days.

As discussed in Rigobon and Sack (2004), α can be estimated by using $\boldsymbol{\Delta\Omega}$ as the instruments for identification. To formalize the instruments, we first define endogenous variables. Let $\boldsymbol{\Delta}\mathbf{i}_A$ and $\boldsymbol{\Delta}\mathbf{s}_A$ be $T_A \times 1$ vectors of variables on the announcement days, and $\boldsymbol{\Delta}\mathbf{i}_{\bar{A}}$ and $\boldsymbol{\Delta}\mathbf{s}_{\bar{A}}$ be $T_{\bar{A}} \times 1$ vectors of variables on the non-announcement days. Then, we can combine these two vectors into $T \times 1$ vectors of endogenous variables:

$$\boldsymbol{\Delta}\mathbf{i} \equiv [\boldsymbol{\Delta}\mathbf{i}_A, \quad \boldsymbol{\Delta}\mathbf{i}_{\bar{A}}]', \tag{1.4}$$

$$\boldsymbol{\Delta}\mathbf{s} \equiv [\boldsymbol{\Delta}\mathbf{s}_A, \quad \boldsymbol{\Delta}\mathbf{s}_{\bar{A}}]'. \tag{1.5}$$

Given these endogenous variables, $\boldsymbol{\Delta}\mathbf{i}$ and $\boldsymbol{\Delta}\mathbf{s}$, instruments are constructed by normalizing with the number of days in each subset of days, and by flipping the signs of the variables on the non-announcement days:

$$\mathbf{z_i} \equiv \left[\frac{1}{T_A}\boldsymbol{\Delta}\mathbf{i}_A, \quad -\frac{1}{T_{\bar{A}}}\boldsymbol{\Delta}\mathbf{i}_{\bar{A}}\right]', \tag{1.6}$$

$$\mathbf{z_s} \equiv \left[\frac{1}{T_A}\boldsymbol{\Delta}\mathbf{s}_A, \quad -\frac{1}{T_{\bar{A}}}\boldsymbol{\Delta}\mathbf{s}_{\bar{A}}\right]'. \tag{1.7}$$

[5] For the derivation, see Appendix A.2.

It is easy to see that $\mathbf{z_i}$ and $\mathbf{z_s}$ are relevant and valid instruments to identify the pass-through in Equation (1.2). First, these instruments are correlated with the endogenous variables as long as the variances on the announcement days and non-announcement days are different. On the other hand, these instruments are uncorrelated with the shocks to the asset prices as presented in Equation (1.3).[6]

In this paper, I use the orthogonality of both instruments as the moment conditions for GMM estimation. Though I can estimate the pass-through by the IV estimation using just one instrument, as implemented in Rigobon and Sack (2004), the GMM estimation should provide more efficient estimates. The moment conditions are described as follows:

$$E[f_t(\alpha)] = 0, \tag{1.8}$$

where

$$\begin{aligned} f_t(\alpha) &= Z_t \cdot e_t, \\ Z_t &= [z_{i,t},\, z_{s,t}]', \\ e_t &= \Delta s_t - \alpha \Delta i_t. \end{aligned}$$

The GMM estimate of α can be obtained by solving the minimum distance problem:

$$\alpha_{GMM} = \arg\min_{\alpha} f_T(\alpha)' \, W \, f_T(\alpha), \tag{1.9}$$

where $f_T(\alpha) = \sum_{t=1}^{T} f_t(\alpha)$ and W is an appropriate 2×2 weighting matrix. I use the two-step GMM for the estimation, in which I first use the identity

[6] For the formal proof of the orthogonality, see Appendix A.3.

matrix as a weighting matrix to solve the minimization problem, and then use the inverse of estimated variance-covariance matrix of the moment conditions in the first step as a weighting matrix in the second step. Inference for the GMM estimation is based on heteroscedasticity-robust standard errors.

1.2.4 Weak-Identification Robust Confidence Set

One concern of identification through heteroscedasticity is that the difference between the announcement days and non-announcement days may not be large enough for strong identification. In order to address the issue of weak identification, I employ the two statistic that could derive weak-instrument robust confidence sets: the S statistic in Stock and Wright (2000), an extension of the Anderson and Rubin's (1949) statistic to GMM, and the K statistic in Kleibergen (2005).[7] These statistics test the null hypothesis for a hypothesized value of the parameter, based on the moment conditions evaluated at the hypothesized value. The confidence set is derived as the set of parameter values for which the test accepts the null hypothesis.

The S statistic is defined as follows:

$$S(\alpha_0) = \left[\sqrt{\frac{1}{T}} f_T(\alpha_0)\right]' \hat{V}_{ff}(\alpha_0)^{-1} \left[\sqrt{\frac{1}{T}} f_T(\alpha_0)\right], \tag{1.10}$$

where

$$\hat{V}_{ff}(\alpha) = Var\left(\sqrt{\frac{1}{T}} f_T(\alpha)\right), \tag{1.11}$$

which is an estimate of the asymptotic variance-covariance matrix of the moment conditions. Note that the S statistic is based on the objective function in the minimization problem in Equation 1.9, but the value is evaluated at the

[7] For details, see Stock, Wright, and Yogo (2002).

hypothesized value of the parameter, α_0. The S statistic has a chi-square null limiting distribution, with the number of moment conditions as the degrees of freedom.

The K statistic is defined as follows:

$$K(\alpha_0) = \frac{1}{4T}\left(\left.\frac{\partial S(\alpha)}{\partial \alpha}\right|_{\alpha_0}\right)[\hat{D}'_T(\alpha_0)\hat{V}_{ff}(\alpha_0)^{-1}\hat{D}_T(\alpha_0)]^{-1}\left(\left.\frac{\partial S(\alpha)}{\partial \alpha}\right|_{\alpha_0}\right)', \tag{1.12}$$

$$= \left[\sqrt{\frac{1}{T}}f_T(\alpha_0)\right]'\left[\hat{V}_{ff}(\alpha_0)^{-\frac{1}{2}}P_{\hat{V}_{ff}(\alpha_0)^{-\frac{1}{2}}\hat{D}_T(\alpha_0)}\hat{V}_{ff}(\alpha_0)^{-\frac{1}{2}}\right]\left[\sqrt{\frac{1}{T}}f_T(\alpha_0)\right], \tag{1.13}$$

where

$$\hat{D}_T(\alpha) = q_T(\alpha) - \hat{V}_{\alpha f}(\alpha)\hat{V}_{ff}(\alpha)^{-1}f_T(\alpha),$$

$$q_T(\alpha) = \frac{\partial f_T(\alpha)}{\partial \alpha},$$

$$\hat{V}_{\alpha f}(\alpha) = Cov\left(\sqrt{\frac{1}{T}}f_T(\alpha), \sqrt{\frac{1}{T}}q_T(\alpha)\right), \quad \text{and}$$

$$P_{\hat{V}_{ff}(\alpha_0)^{-\frac{1}{2}}\hat{D}_T(\alpha_0)} = \hat{V}_{ff}(\alpha_0)^{-\frac{1}{2}}\hat{D}_T(\alpha_0)[\hat{D}_T(\alpha_0)'\hat{V}_{ff}(\alpha_0)^{-1}\hat{D}_T(\alpha_0)]^{-1}\hat{D}_T(\alpha_0)'\hat{V}_{ff}(\alpha_0)^{-\frac{1}{2}}.$$

Essentially, the K statistic uses an optimal subset of moment conditions to improve the power of the tests. In other words, by using a subset of more relevant moment conditions, we could improve the efficiency of the test statistic, which leads to the higher power of the tests. The only difference between the S statistic in Equation (1.10) and the K statistic in Equation (1.12) is that the K statistic uses the variance-covariance matrix adjusted by the projection matrix based on

$\hat{D}_T(\alpha)$. $\hat{D}_T(\alpha)$ is a residual of the gradient of moment conditions, $q_T(\alpha)$, after projecting it on the level of the moment conditions, $f_T(\alpha)$. By construction, $\hat{D}_T(\alpha)$ is orthogonal to $f_T(\alpha)$ and we use this orthogonality to improve the efficiency. The K statistic also has a chi-square null limiting distribution, with the number of parameters as the degrees of freedom.

1.3 Data and Background

This section describes the data and background to show that there is enough variation to identify the pass-through of monetary policy shock. After describing the data, I provide a brief summary of Japanese monetary policy. Then, I explain the selection of important announcements and provide a statistical analysis showing that the selection is valid in terms of identification.

1.3.1 Data

I estimate the pass-through of monetary policy shocks to three financial assets: (1) corporate bond yields, (2) stock prices, and (3) the exchange rate. The analysis is based on daily data from April 1998 to July 2013, all of which are obtained from the Bloomberg.

The details of the series are described as follows:

1. Japanese Government Bond (JGB) yield: generic yield with a maturity of 5, 7, 10 and 20 years;

2. Corporate bond yield: the Bloomberg Fair Value (BFV) indices of AA and BBB corporate bond yields for the industrial sector, with a maturity

of 5 and 10 years;[8]

3. Stock prices: Nikkei 225 and TOPIX (stock price index in the Tokyo stock exchange [TSE]);

4. Exchange rates: spot exchange rates of the U.S. dollar and the euro, measured by the Japanese yen.

I compute the daily changes in levels for the JGB and corporate bond yields, and the continuously compounding rate of daily change for stock prices and the exchange rate. The data between March 11 and March 18, 2011 is excluded from the analysis to eliminate the effect of the earthquake in March 2011.

1.3.2 Brief Summary of Japanese Monetary Policy

The BOJ's overnight interest rate has been reduced to nearly zero for approximately two decades, and three different programs of unconventional monetary policies have been implemented. Table 1.1 summarizes the timeline of the important events in Japanese monetary policy, and Figure 1.2 shows the volume of monetary base.[9]

The first program was called "quantitative easing" (QE) and implemented from March 2001 to March 2006 under Governors Hayami and Fukui. Under the QE program, the BOJ set its current account balance as the main policy target and purchased long-term JGBs to achieve this target.[10] The QE program

[8] The AA index and the BBB index are available from June 8, 1999 and January 30, 2003, respectively. The BBB index with a maturity of 10 years is not available from February 6, 2012.

[9] For more comprehensive accounts of BOJ's monetary policy, see Ito and Mishkin (2006) and Ueda (2012b).

[10] For an analysis of the QE program, see Ugai (2007) and Shiratsuka (2010). More recently, Shibamoto and Tachibana (2013) analyze the effects of the QE program using the method of

was terminated in March 2006, and the overnight interest rate was gradually raised to 0.5 percent. However, in response to the global financial crisis in 2008, the BOJ reduced the overnight interest rate to zero again.

The second program was called "comprehensive monetary easing" (CME) and implemented from October 2010 to April 2013 under Governor Shirakawa. Under the CME program, the BOJ purchased not only JGBs, but also commercial paper and risky assets such as ETFs and REITs, while also making a commitment to keep the policy rate at zero. This is the policy prescription proposed by Eggertsson and Woodford (2003). In addition, the BOJ provided various forms of lending programs to financial institutions.

Most recently, the BOJ substantially expanded its asset purchase program and launched it as "quantitative and qualitative monetary easing" (QQME). The QQME program has been implemented since April 2013 under Governor Kuroda. Under the QQME program, the BOJ commits to achieving 2 percent inflation by 2015. To achieve this goal, the BOJ announced that it would increase the monetary base at a pace of 60 to 70 trillion yen per year, and extend the average maturity of its JGB holdings from three years to seven years.[11]

1.3.3 Selection of Monetary Policy Announcements

The selection of the important announcements is crucial for the analysis in this paper, and I select 41 announcement days from April 1998 to July 2013 based

identification through heteroscedasticity. Kimura and Nakajima (2013) use a special type of regime-switching structural VAR using the data from 1981 to 2012, with ad hoc shrinkage in certain parameters.

[11] For details, see Kuroda (2013).

on Ueda (2012a). The exact dates and overview are listed on Table 1.2. These dates are associated with the BOJ's "official" change in its monetary policy.[12]

In addition, I include other dates when strong signals concerning future changes in BOJ's monetary policy were made. For example, I include the days when the new BOJ governor was nominated and the confirmation hearing was held by the National Diet in 2013. On the other hand, I do not include other meeting days or the days of speeches by the BOJ governor or other board members. This is because including trivial or indirect news announcement days will make the distinction between the announcement days and the non-announcement days unclear, which undermines the identification as discussed in Wright (2012).

1.3.4 Standard Deviations of the Series

To see whether the actual data is consistent with the assumptions for the identification, Table 1.3 compares the standard deviations of the daily changes on the announcement days and non-announcement days. Consistent with the assumptions, the standard deviations are higher on the announcement days than on the non-announcement days for almost all series, except for the BBB yield with a maturity of 10 years.

To test whether the variances in these two sets of days are significantly different, I conduct three statistical tests: the F-test, the block bootstrap and the

[12] I exclude two dates from the list in Ueda (2012a), April 13, 1999 and March 14, 2011. First, I exclude April 13, 1999 because the commitment to keep the policy rate at zero made by the BOJ governor is not entirely clear. Ueda (2012a) also notes that the market reacted to this event very slowly. However, including this date does not materially change the results. Second, I exclude March 14, 2011 since this is the meeting right after the earthquake.

stationary bootstrap. All tests are based on the null hypothesis that the population variance in these two sets of days are equal. Since the F-test assumes that each observation is independent, I use the block bootstrap and the stationary bootstrap to take the heteroscedasticity observed in the daily data into account. In the block bootstrap, I construct an artificial sample by resampling the block of 10 days to preserve heteroscedasticity. On the other hand, in the stationary bootstrap, the length of the block is randomly determined.[13] After constructing an artificial sample, I compute the variance ratio in the artificial sample. By repeating this exercise, I can form the bootstrap distribution of the variance ratio and compute the p-values based on the percentile of the sample variance ratio.

The null hypothesis is rejected for most series, with the sample variance ratio larger than one, except for the 10-year BBB corporate bond yields. Therefore, we could conclude that the variance is significantly larger on the announcement days. The F-test tends to reject the null hypothesis more often than the other two tests based on the bootstrap.

1.4 Results

In this section, I present GMM estimates showing that the pass-through of monetary policy shock is one to one for corporate bond yields, but it is smaller than the U.S. estimates for stock prices and the exchange rate. Then, I discuss how the segmentation of financial markets and the length of the period at the ZLB can help reconcile such differences. In addition, I provide an event

[13] More specifically, an artificial sample is constructed by the Bernoulli trial, either to pick a random sample or the sample on the next day. I set the probability of the latter as 0.9 to make the expected length of the block 10 days.

study showing that the announcements in 2013 have substantial effects on asset prices. Given these results, I discuss the decomposition of the effect of unconventional monetary policy. The results of weak-identification robust confidence set suggest that focusing on the subset of more important announcement will help identification. Lastly, I present several robustness checks using additional variables and subperiods to confirm the main results.

1.4.1 Event Study

Table 1.4 presents the results of the event study focusing on the QQME program in 2013.[14] The results show that the announcements of the QQME program led to a substantial decline in the long-term JGB and corporate bond yields. Furthermore, stock prices substantially increased on these announcement days, but the effect on exchange rates was mixed.

The results show that the long-term JGB and corporate bond yields substantially declined responding to the announcements. Specifically, the JGB yields declined 11.4 (10 years) and 17.7 (20 years) basis points on April 4th, and the cumulative decline on all announcements was 14.0 (10 years) and 23.3 (20 years) basis points, respectively. Corporate bond yields also declined on these days, but the magnitude is smaller than JGB yields. The 10-year AA bond yield declined 9.72 basis points on April 4, and the cumulative decline was 10.65 basis points.[15]

[14] Ueda (2012a) provides an event study of the monetary policy announcements between 1999 and 2011 and shows that the announcements regarding the QE programs lowered the JGB and corporate bond yields, but did not significantly affect stock prices or exchange rates. Lam (2011) conducts an event study from 2008 to 2011, which covers the part of the CME program, and finds similar results.

[15] The effects for yields with shorter maturities are trivial, primarily because they are all stuck at the zero lower bound.

The results also show that stock prices substantially increased on the announcement days. The cumulative increases of the Nikkei 225 and TOPIX indices were 5.18 percent and 3.68 percent, respectively. On the other hand, the effect on exchange rates was mixed. Even though the announcement on April 4 led to a substantial depreciation of the Japanese yen, 3.49 percent for the U.S. dollar and 4.16 percent for the euro, this effect was quickly offset on subsequent announcements.

1.4.2 GMM Estimates of Pass-Through

Table 1.5 presents the pass-through to corporate bond yields, stock prices, and the exchange rate. The results show that there is a statistically significant and nearly one-for-one pass-through to corporate bond yields, notably for high-grade bond yields. On the other hand, the pass-through to stock prices and the exchange rate is negative but the estimated magnitude is quite small and not statistically significant in most cases.

I present the estimates of the pass-through to different asset prices based on an individual JGB yield. This is because monetary policy shocks primarily influence the overall level of the JGB yields, and the change in JGB yield with any maturity should reflect such shocks.

1.4.2.1 Corporate Bond Yields

For the AA corporate bond yields, most of the pass-through is statistically significant, with a magnitude from 0.39 to 1.84. It implies that an expansionary monetary policy shock, which lowers the 20-year JGB yield by 100 basis points, will lower the 5-year AA corporate bond yield 39 basis points, and lower the 10-year AA corporate bond yield 92 basis points. All estimates are smaller than one and statistically significant, except for the pass-through from the 5-year JGB to the 10-year AA bond.

On the other hand, the estimates of the pass-through to the medium-grade corporate bond yields varies; the estimated magnitude is from -2.90 to 2.50. The estimates are positive in most cases and some of them are close to one, which suggests that monetary policy shocks are passed on to BBB corporate bond yields to some extent. However, no estimate is statistically significant because of the large standard errors.

1.4.2.2 Stock Prices

The estimates of the pass-through to stock prices are mostly negative, ranging between -1.32 and -0.13. These estimates imply that an expansionary monetary policy shock, which lowers JGB yields by 100 basis points, increases stock prices by 0.13 to 1.32 percent. However, only one of them, the pass-through from the 20-year JGB yield to the TOPIX index is statistically significant. In other words, the estimated magnitude of the pass-through to stock prices is so small that the estimates are not significantly different from zero in most cases.

1.4.2.3 Exchange Rates

The results show that the pass-through to the exchange rate is mostly negative, with the magnitude between -0.77 and 0.09. These estimates imply that an expansionary monetary policy shock, which lowers JGB yields by 100 basis points, leads a depreciation of the Japanese yen by 0.77 percent, or its appreciation by

0.09 percent.[16] However, only the pass-through from the 20-year JGB yield, -0.21 for the U.S. dollar and -0.20 for the euro, is statistically significant.

These results are consistent with the prediction of conventional interest rate parity, which suggests that the expected return on domestic assets should be the same as the exchange-rate adjusted expected return on foreign assets. In other words, a decline in the JGB yield should be adjusted by a depreciation of the Japanese yen, which would boost the stimulative effect of expansionary monetary policy shocks. However, the estimated magnitude of the pass-through is quite small and few estimates are statistically significant.

1.4.3 Weak-Identification Robust Confidence Set

As can be seen from the results, the S and the K tests fail to identify the relevant confidence set in many cases. In other words, the tests accept the null hypothesis for all parameter values. This lack of identification suggests that the difference of the variances on the announcement and non-announcement days is not big enough for identification, even though the difference is statistically significant in some tests.

To show that focusing on more relevant announcements improves identification, Table 1.6 presents the results with 25 selected announcements that are explicitly related to JGB purchases and the QE programs. In contrast to the results based on the baseline set of announcements, both the S and K tests identify the relevant confidence sets in many cases, which suggests that eliminating less relevant announcements greatly improves identification. The estimates are

[16] Since I use the exchange rate that measures the value of non-Japanese currency by the Japanese yen, a rise in the exchange rate implies depreciation of the Japanese yen.

broadly similar to the results based on the baseline results. The pass-through to corporate bond yields is statistically significant in most cases and its magnitude is close to one. Unlike the baseline case, the pass-through to stock prices is often positive, but its magnitude is small and not statistically significant. The pass-through to the exchange rate is mostly negative, and some estimates are statistically significant.

1.4.4 Discussion of the Results

Given these results, I first provide the comparison between Japanese and the U.S. results. Though Japanese results are consistent with the U.S. results in their signs, the estimated magnitude of pass-through is considerably smaller for stock prices and the exchange rate. I discuss that the higher degree of market segmentation, and the smaller scope to achieve further accommodation using forward guidance could help explain such differences. Second, I compare the event study focusing on announcements in 2013 and the GMM estimates using the entire sample to highlight that the announcements in 2013 had a substantial effect on asset prices. Then I associate this difference with the regime change of BOJ's monetary policy. Lastly, I make a conjecture about the decomposition of the effect of unconventional monetary policy based on survey measures, to argue that it primarily affects the term premium.

1.4.4.1 Comparison with the U.S. Estimates

For corporate bond yields, Japanese and U.S. estimates are similar both in their signs and magnitudes. On the other hand, though both the Japanese and the U.S. estimates are negative for stock prices and the exchange rate, which is consistent with the predictions of economic theories, the magnitude of the Japanese estimates is substantially smaller than the U.S. estimates.

For corporate bond yields, the pass-through to corporate bond yields is broadly similar in Japan and in the U.S. Raskin (2013a) finds the one-to-one pass-through from the U.S. treasury yields to high-grade corporate bond yields with the ratings above A, based on identification through heteroscedasticity. Similar to the Japanese case, the pass-through to medium-grade bond yields with the ratings of BBB and BB is statistically insignificant in most cases.[17]

For stock prices, the pass-through is negative both in Japan and the U.S., but the estimated magnitudes are substantially smaller in Japan. Kiley (2013) estimates that the responses of the stock prices to monetary policy shocks ranges between -3.0 and -1.5, by using the sample in the Great Recession.[18] Compared to these estimates, the Japanese estimates, which are marginally negative with a magnitude of 0.2, are substantially smaller.[19]

[17] Krishnamurthy and Vissing-Jorgensen (2011) also note that the LSAP had a negligible effect on the medium-grade bond yields.These results are puzzling because corporate bond yields could be regarded as the sum of the risk-free rate and the risk premium for a default. If an expansionary monetary shock lowers the government bond yields, it should equally affect the high-grade and medium-grade corporate bond yields. One way to explain the limited pass-through to the medium-grade bond yields is to introduce the safety premium proposed by Krishnamurthy and Vissing-Jorgensen (2012). The safety premium is the premium paid by investors to satisfy their unique demand for long-term safe assets. Since the medium-grade corporate bonds are not regarded as safe assets, their yield are not affected by the decline of the government bond yields. Krishnamurthy and Vissing-Jorgensen (2011) argue that the effect of the Fed's LSAP announcements is primarily due to reduction in the safety premium.

[18] Kiley (2013)'s estimates are substantially smaller than the estimates in Gürkaynak, Sack, and Swanson (2005) and Bernanke and Kuttner (2005) that use the sample before the Great Recession. This is because monetary policy shocks identified in the ZLB sample are less stimulative than in normal times due to the lower bound on short-term yields. In other words, the effect of monetary policy shocks may be weaker at the ZLB because they can only influence a part of the yield curve.

[19] Similar to Japanese case, Rosa (2012) also documents that the effect of Bank of England's announcements about gilt purchases is not statistically significant on stock prices.

For the exchange rate, the magnitude of Japanese estimates, -0.2, is also substantially smaller than the U.S. estimates, even though the sign of pass-through is the same in both countries. Neely (2013) estimates that the announcement of the Fed's LSAP caused a 3 or 4 percent depreciation of the U.S. dollar. Glick and Leduc (2013) also report that the pass-through to exchange rates during the Great Recession is -3.0 by identifying the monetary policy surprises based on high-frequency data.

To interpret the smaller magnitudes of the pass-through in Japan compared to the U.S., I propose two possible explanations: First, the Japanese stock and foreign exchange markets may not be responsive to monetary policy shocks because of the higher degree of market segmentation. For example, it is common for Japanese firms to hold each other's stocks to prevent hostile takeovers. On the other hand, there are retail investors in the foreign exchange market, who constitute a considerable portion of the market and tend to hold foreign-currency dominated assets for a long period.[20] Since these investors wouldn't change their allocation of assets responding to monetary policy shocks, the reactions of asset prices become less pronounced and the overall pass-through could be smaller. In other words, these institutional factors increase the degree of segmentation in the stock and foreign exchange markets, which may lead to inefficient monetary transmission mechanisms.

Second, the scope to provide further accommodation using forward guidance may be smaller in Japan, because the timing of any expected liftoff is far away. Since the Japanese economy has been stuck at the ZLB for nearly two decades, it is very hard for any announcement to change expectations about future short

[20] For details, see Zurawski and D'Arcy (2009).

rates. As can be seen in Figure 1.3, the JGB yield curve has shifted downward over time, and even the 5-year JGB yield is virtually zero after 2010. As Swanson and Williams (forthcoming) pointed out, asset prices are affected by the entire path of expected future short-term interest rates, not only the current level of interest rate. However, this extremely low level of the long-term JGB yields suggest that the expected lift off of short-term interest rate from zero is far away, which makes the effect of forward guidance policy quite limited. As a result, the announcements cannot affect the asset prices and the pass-through of monetary policy shocks becomes smaller.

1.4.4.2 Comparison of Announcements Before and After 2013

The comparison of the event study and GMM show that the announcements in 2013 had substantial effects, not only on corporate bonds, but also on stock prices. It may be because the announcements in 2013 were associated with the BOJ's regime change, in which the BOJ made an open-ended commitment to achieve the inflation target, creating a huge surprise.

Romer (2013) describes this policy change in 2013 as "an honest-to-goodness regime shift" and "(the BOJ) took dramatic actions and pledged convincingly to do whatever it takes to end deflation." Consistent with her argument, some survey measures show that inflation expectations have substantially increased in 2013. For example, Figure 1.4 shows the upward shift of the inflation expectation in the mean of the ESP Forecast (survey of professional forecasters in Japan) in 2013, which could be up to 0.8 percent. Hausman and Wieland (2014) also provide similar analysis and conclude that the set of policy package called "Abenomics," which includes the introduction of the QQME program in 2013, raised long-run inflation expectation.[21] However, even though the announcements in 2013 have substantial effects, the GMM estimates are not statistically significant since they average the effect over the full sample.

1.4.4.3 Decomposition into Different Component and Individual Policy

Lastly, I provide the discussion of the decomposition of the effects of unconventional monetary policy. I discuss two types of decompositions; the first is to analyze which component of long-term interest rate is affected by unconventional monetary policy, and the second is to analyze the effects of different policies, either forward guidance or asset purchases.

For the first decomposition, I make a conjecture that unconventional monetary policy primarily affects the term premium, based on the survey measures. As summarized in Bernanke (2013), the long-term interest rate can be decomposed into three components (the expected path of inflation, the expected path of the real interest rate, and the term premium), and a growing body of research claims that the Fed's LSAP primarily affects the term premium.

In Japan, it seems likely that the QE programs had primarily influenced the term premium, because survey measures of inflation expectations had been quite stable until recently, and so had survey measures of short-term interest rates. Furthermore, considering that the long-term yields declined after the introduction of QQME program, despite the increase in the inflation expectations, it is even more likely that these announcements reduced the term premium. However, it is hard to quantify the exact effects due to the lack of a credible measure

[21] On the other hand, Fujiwara, Nakazono, and Ueda (2014) note that there is no significant increase in inflation expectation at 10-year horizon after the introduction of the QQME.

of inflation expectations in Japan.[22]

For the second decomposition, this analysis only provides the estimates of the *average* effects of unconventional policies, and estimates cannot be interpreted as the effect of individual policies, forward guidance or asset purchases. In the recent literature, the channels through which QE programs have an impact on financial markets become very topical.[23] I provide the analysis of subsample to shed some light on this issue, but it is difficult to tackle this decomposition due to the simultaneity of the announcements and lack of high-frequency data, and I will leave it as a future research question.

1.4.5 Robustness Checks

In this section, I provide several robustness checks, to which the main results are generally robust. First, I analyze the pass-through to other financial assets. Second, I consider subsamples focusing on different programs. Third, I provide the analysis based on the principal component of JGB yields. Last, I use alternative sets of non-announcement days.

1.4.5.1 Additional Variables

I extend the analysis using three kinds of additional variables: (1) REITs, (2) CDS index, and (3) exchange rates in other OECD and Asian economies.[24]

[22] For example, the Japanese break-even inflation rate is not reliable as a measure of inflation expectation since its market is too small and the issuance of the new inflation-index linked bonds has been suspended since October 2008.

[23] For example, Woodford (2012) points out that distinguishing the effect of the central bank's economic outlook and the effect of policy announcements is crucial. Krishnamurthy and Vissing-Jorgensen (2013) claim that evaluation of the channels of an unconventional monetary policy is necessary when considering the exit strategy. Some papers such as Raskin (2013b) explicitly analyze the effect of forward guidance policy.

[24] The REIT index is available from April 2003.

These variables are of interest for slightly different reasons. First, REITs are of interest because the BOJ purchased them in the QE programs. Second, the analysis of the CDS index will shed some light on the effects of the QE programs on credit default risk as discussed in Gilchrist and Zakrajsek (2013). Lastly, I focus on exchange rates relative to these countries because the so-called "yen-carry trade," in which investors borrow money in the Japanese yen and invest in high-interest rate currencies, has been prominent since the late 1990s. These high-yield currencies include the Australian and New Zealand dollars and other emerging Asian currencies.[25]

Table 1.7 presents the standard deviations and variance ratios of these variables. Consistent with the assumptions of the identification, the standard deviations are higher on the announcement days than on the non-announcement days for all series. In addition, most of these differences are statistically significant.

Table 1.8 shows the GMM estimates of the pass-through to the REITs, the CDS index, and exchange rates using the baseline announcement days, and Table 1.9 shows the GMM estimates using the selected announcement days. For the REITs, the pass-through is mostly negative and some of them are statistically significant. The negative pass-through implies that an expansionary monetary policy shock lead to the increase in the price of these financial assets. The magnitude of the pass-through to the REITs is between -0.46 and -0.16. The pass-through to the CDS index is positive, but the magnitude varies substantially with the maturity of the JGB yields and none are statistically significant. The pass-through to exchange rates in the OECD and Asian economies is negative in almost all cases, with a magnitude between -0.14 and -0.22. The

[25] For details, see Hattori and Shin (2008).

estimates are similar to the estimates for the U.S. dollar and the euro, which suggests that the expansionary monetary policy shocks lead a depreciation of the Japanese yen relative to all other currencies. However, the magnitude of the depreciation is quite small.

1.4.5.2 Analysis of Subsamples

To analyze how different QE programs affect corporate bond yields and asset prices, I conduct the same analysis in three subsamples focusing on the different QE programs: (1) the QE program (2001-2006), (2) the period between the QE and the CME programs (2006-2010), and (3) the CME program (2010-2013). The results are presented in Table 1.10. The period before the QE program (1998-2001) and the QQME program (2013-) are excluded since the number of announcements are too small.

The findings are broadly similar to the main results. The QE programs have statistically significant pass-through to corporate bond yields, mostly one-to-one, but the pass-through to the stock prices or exchange rates is not statistically significant. For the high-grade corporate bond yields, pass-through is mostly one-to-one with a magnitude between 0.36 and 1.87. Unlike the main results, the pass-through to the medium-grade bond yields is statistically significant in some cases, with the magnitude between 0.44 and 1.98. Such a wide range of estimates may reflect their volatility due to small samples in the subperiods. The signs of pass-through to stock prices are mixed but none are statistically significant. The pass-through to exchange rates is negative in most cases, but only a few estimates are statistically significant.

Furthermore, there is no obvious difference in the pass-through in different subperiods, which suggests that the effects of different QE programs are broadly similar over time. One evident difference is that the pass-through from the 5-year JGB yield is not statistically significant during the CME program between 2010 and 2013. But this is because the 5-year JGB yield is extremely low in this period.

1.4.5.3 Analysis Using the Principal Component of JGB Yields

To jointly analyze the pass-through from the JGB yields with different maturities, I conduct the same analysis using a principal component of these JGB yields. Table 1.11 presents the results for main and additional variables. The results are broadly similar to the main results. The pass-through to the high-grade corporate bond yields is positive and statistically significant, whereas the pass-through to the stock prices and exchange rates is negative but not statistically significant in most cases.

1.4.5.4 Alternative Set of Non-Announcement Days

I use two alternative definitions of non-announcement days to estimate the pass-through of monetary policy shock. First, following Rigobon and Sack (2004), a non-announcement day is defined as one business day prior to the announcement day. Accordingly, the number of announcement days and non-announcement days are the same. Second, following Gilchrist and Zakrajsek (2013), I exclude all the BOJ meeting days from the set of non-announcement days because some trivial or indirect news released on these meeting days may contaminate the identification. The estimates using these alternative definitions are not listed to conserve space, but the results are essentially the same as the baseline set of non-announcement days.

1.5 Conclusion

This paper analyzes the effects of unconventional monetary policy in Japan, which has been stuck at the ZLB for a substantially longer period than any other economy. To discuss if we could learn any lessons from Japanese experience, this paper compares the estimates of the pass-through of monetary policy shocks with the U.S. estimates.

By using the tools of an event study and identification through heteroscedasticity, I find that the effects of expansionary monetary policy shocks are directly passed on to corporate bond yields, notably for high-grade bond yields. However, the pass-through to stock prices and the exchange rate is not statistically significant in most cases. These results contrast with the U.S. results, where the pass-through to all assets is statistically significant. Such differences may reflect the segmentation of Japanese financial markets, or the limited scope for the forward guidance policy to be effective in Japan, due to the prolonged period at the ZLB.

The comparison between Japan and the U.S. suggests two important implications for policymakers: First, unconventional monetary policy is effective even in an extreme environment where the economy is stuck at the ZLB for nearly two decades. Second, however, its effect could be substantially limited if the financial market is segmented, or the economy is stuck at the ZLB for a prolonged period.

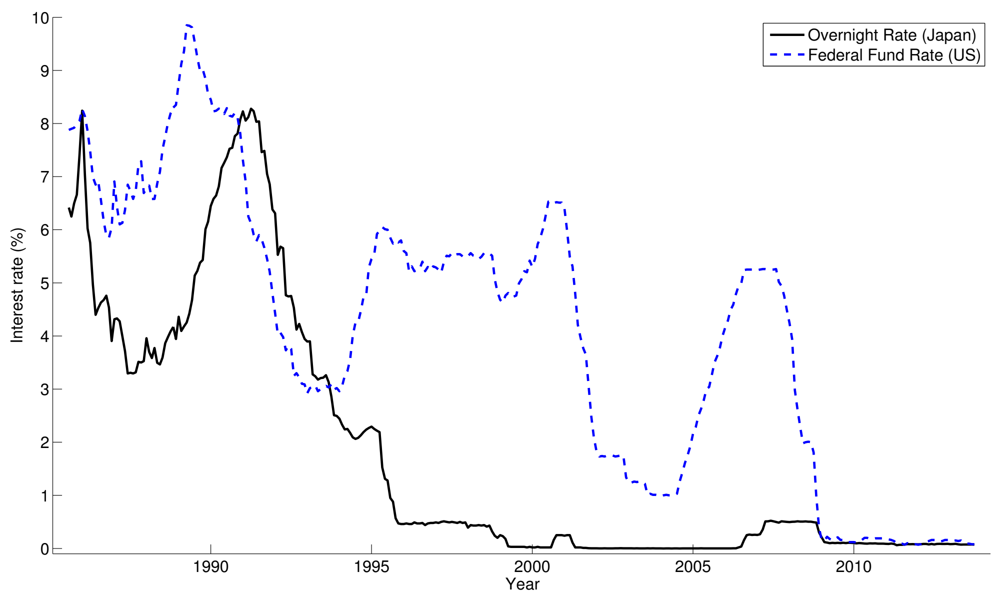

Figure 1.1: BOJ's overnight policy rate and Fed's federal fund rate from 1985 to the present

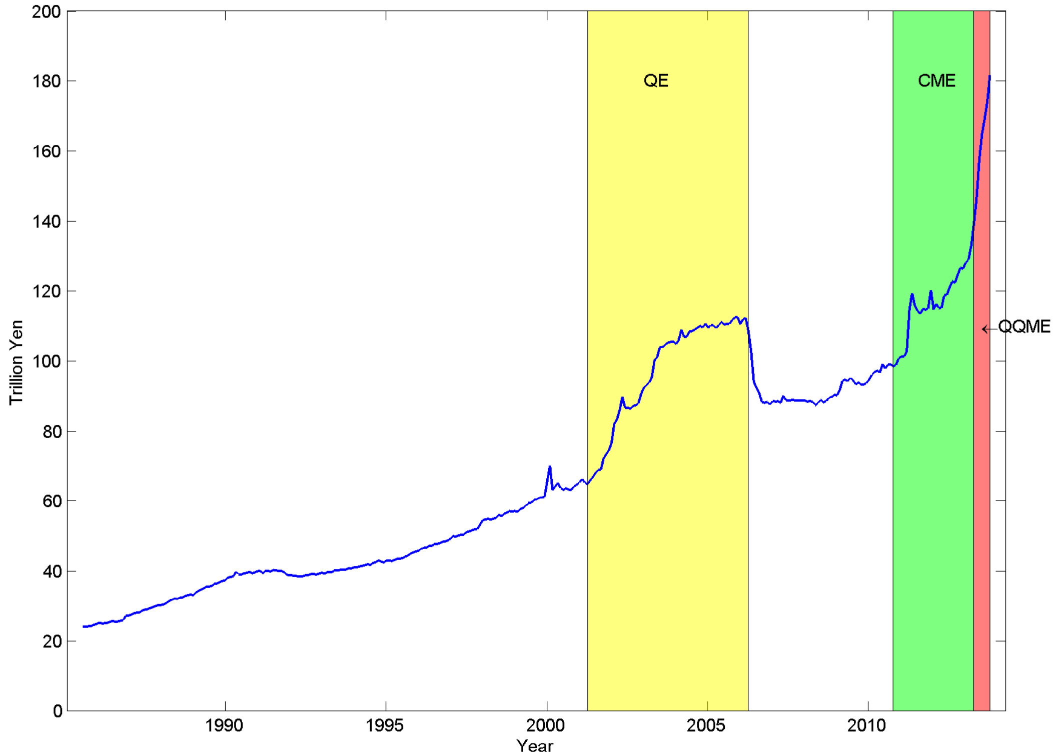

Figure 1.2: Monetary Base from 1985 to the present

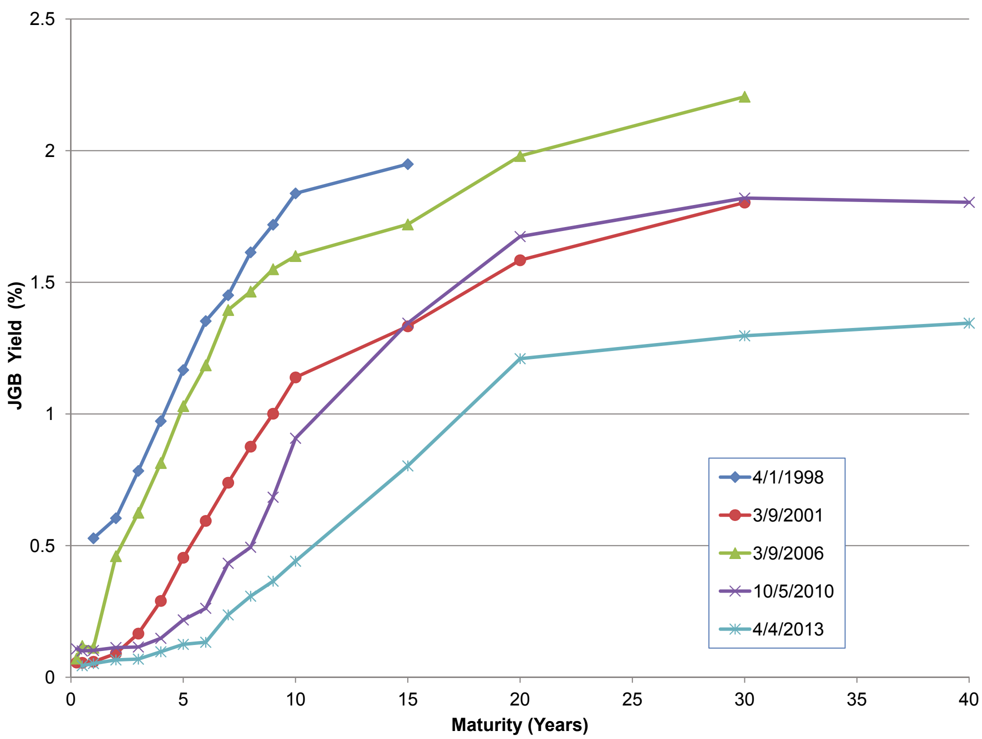

Figure 1.3: JGB Yield Curve from 1998 to 2013 (Selected Dates)

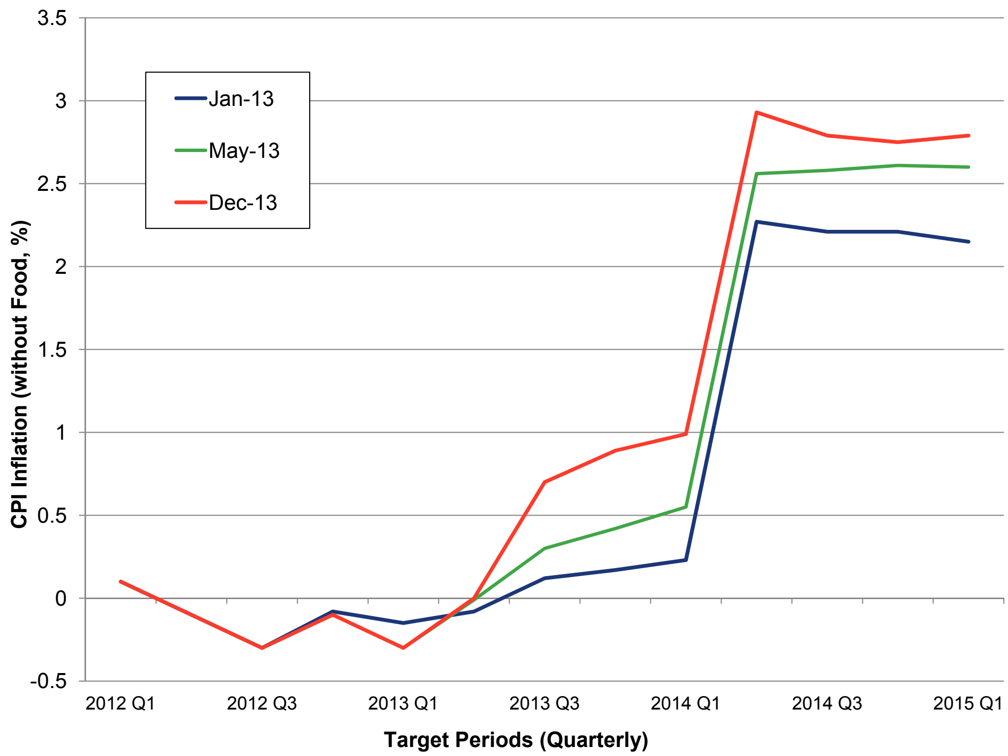

Figure 1.4: ESP Inflation Forecast in 2013 (Target Periods of 2013–2015)

Date	Event	Policy Rate(%)	Governer
4/14/1995		1.0	Matsushita
9/8/1995		0.5	
4/1/1998	The revised BOJ act came in effect	0.5	Hayami
2/12/1999		0	
8/11/2000		0.25	
3/19/2001	Quantitative Easing (QE) launched	0	
3/9/2006	QE terminated	0	Fukui
7/14/2006		0.25	
2/21/2007		0.5	
10/31/2008		0.3	Shirakawa
12/19/2008	JGB purchase increased	0	
10/5/2010	Comprehensive Monetary Easing (CME) launched	0	
4/4/2013	CME terminated, and Quantitative and Qualitative Monetary Easing (QQME) launched	0	Kuroda

Table 1.1: Timeline of BOJ's Monetary Policy

Date	Event	Summary
9/9/1998	BOJ Statement	Policy rate reduced to 0.25 percent
2/12/1999	*BOJ Statement*	Policy rate reduced close to zero
8/11/2000	BOJ Statement	Policy rate raised to 0.25 percent
3/19/2001*	*BOJ Statement*	*Quantitative easing (QE) launched (policy rate close to zero)*
8/14/2001*	*BOJ Statement*	*QE expanded*
12/19/2001*	*BOJ Statement*	*QE expanded*
2/28/2002*	*BOJ Statement*	*JGB purchase increased*
9/18/2002*	BOJ Statement	Stock purchase announced
10/30/2002*	*BOJ Statement*	*QE expanded*
3/25/2003*	BOJ Statement	Stock purchase expanded
4/8/2003*	BOJ Statement	ABS purchase announced
4/30/2003*	*BOJ Statement*	*QE expanded*
5/20/2003*	*BOJ Statement*	*QE expanded*
10/10/2003*	*BOJ Statement*	*QE expanded*
1/20/2004*	*BOJ Statement*	*QE expanded*
3/9/2006*	*BOJ Statement*	*QE terminated*
7/14/2006	BOJ Statement	Policy rate raised to 0.25 percent
2/21/2007	BOJ Statement	Policy rate raised to 0.5 percent
9/18/2008*	BOJ Statement	Dollar swap
10/31/2008	BOJ Statement	Policy rate reduced to 0.3 percent
12/2/2008*	BOJ Statement	Facilitation of corporate finance
12/19/2008*	*BOJ Statement*	*JGB purchase increased (policy rate reduced close to zero)*
2/3/2009*	BOJ Statement	Stock purchase restarted
3/18/2009*	*BOJ Statement*	*JGB purchase increased*
12/1/2009*	BOJ Statement	Fixed-rate 3-month operation
12/18/2009*	BOJ Statement	"Inflation target" clarified
3/17/2010*	BOJ Statement	Fixed-rate operation expanded
5/21/2010*	BOJ Statement	Growth enhancing operation
8/30/2010*	BOJ Statement	Fixed-rate 6-month operation
10/5/2010*	*BOJ Statement*	*Comprehensive monetary easing (CME) launched*
8/4/2011	*BOJ Statement*	*CME expanded*
10/27/2011	*BOJ Statement*	*CME expanded*
2/14/2012	*BOJ Statement*	*CME expanded*
4/10/2012	*BOJ Statement*	*CME expanded*
4/27/2012	*BOJ Statement*	*CME expanded*
10/30/2012	*BOJ Statement*	*CME expanded and the joint statement with the government issued*
12/20/2012	*BOJ Statement*	*CME expanded*
1/22/2013	*BOJ Statement*	*CME expanded and inflation target clarified*
2/25/2013	*Nomination of the new governor*	
3/4/2013	*Confirmation hearing from the new governor at the National Diet*	
4/4/2013	*BOJ Statement*	*Quantitative and qualitative monetary easing (QQME) launched*

a. Dates with superscript are listed in Table 2 of Ueda (2012a). The announcement days that are directly related to the JGB purchase and the QE programs are written in italics.

Table 1.2: Dates of BOJ's Monetary Policy Announcements

Series	Announcement	Non-Announcement	Variance Ratio
JGB			
5 year	3.04	2.89	1.11
7 year	4.47	3.61	$1.53^{**}_{\dagger}$
10 year	4.30	3.33	$1.67^{**}_{\dagger}$
20 year	4.94	3.60	$1.88^{**}_{\dagger}$
Corporate Bond Yield			
AA, 5 year	2.95	2.70	1.19
AA, 10 year	3.88	2.90	$1.79^{**}_{\ddagger}$
BBB, 5 year	2.80	2.78	1.01
BBB, 10 year	3.01	3.12	0.93
Stock Prices			
Nikkei 225	1.98	1.55	$1.64^{**}_{\dagger}$
TOPIX	1.65	1.40	1.39^{*}
Exchange Rates			
US Dollar	1.07	0.71	$2.27^{**}_{\ddagger}$
Euro	1.26	0.85	$2.20^{**}_{\ddagger}$

a. This table compares the standard deviation of daily changes on the announcement days and non-announcement days. Daily changes of the yield in basis points and daily percent changes of the stock prices and exchange rates are used.

b. Superscripts $*, **$ denote the significance at the level of 10% and 5%, respectively, based on the F-test.

c. Subscripts $\dagger, \ddagger$ denote the significance at the level of 10% and 5%, respectively, based on the bootstrap with 10,000 replications. The block bootstrap and the stationary bootstrap lead exactly the same results.

Table 1.3: Standard Deviations and Variance Ratio on the Announcement Days and the Non-Announcement Days

Date	Event	JGB 5 Year	JGB 7 Year	JGB 10 Year	JGB 20 Year	AA 5 Year	AA 10 Year	BBB 5 Year	Nikkei 225	TOPIX	US Dollar	Euro
2/25/2013	*Nomination*	-1.30	-2.00	-2.20	-2.00	-1.37	-1.71	-0.73	2.40	1.77	-1.73	-2.71
3/4/2013	*Confirmation Hearing*	-1.90	-2.30	-4.40	-7.30	-0.69	0.88	-0.96	0.30	0.11	0.79	1.87
4/4/2013	*BOJ Statement*	-1.20	-5.50	-11.40	-17.70	-2.08	-9.72	-1.65	2.18	2.67	3.49	4.16
4/26/2013	BOJ Statement	1.50	1.40	0.80	-0.70	3.02	-0.71	1.83	-0.30	-0.99	-1.23	-1.10
5/22/2013	BOJ Statement	-1.20	0.50	1.20	1.70	-1.04	1.27	-1.02	1.59	0.44	0.64	0.26
6/11/2013	BOJ Statement	4.20	6.30	4.70	4.50	5.39	5.57	5.19	-1.47	-0.98	-2.80	-2.39
7/11/2013	BOJ Statement	-0.80	-3.80	-2.70	-1.80	-3.94	-2.61	-2.06	0.39	-0.04	-0.72	0.20
Baseline Events		**-4.40**	**-9.80**	**-18.00**	**-27.00**	**-4.14**	**-14.17**	**-3.76**	**4.97**	**5.24**	**1.64**	**1.37**
Cumulative		**-0.70**	**-5.40**	**-14.00**	**-23.30**	**-0.71**	**-10.65**	**0.18**	**5.18**	**3.68**	**-2.47**	**-1.65**

a. This table shows the daily changes of the yields and the daily percent changes of stock prices and exchange rates on the days listed above.
b. Baseline events are listed in Table 1.2 and written in *Italic*.

Table 1.4: Effects of the QQME Announcements on JGB, Corporate Bond Yields, Stock Prices and Exchange Rates

	Estimate	Std. Error	S Conf.	K Conf.
Panel A: Pass-Through from 5-Year JGB				
AA, 5 year	0.97**	(0.29)		[0.91, 1.12]
AA, 10 year	1.84**	(0.74)		
BBB, 5 year	1.19	(2.10)		
BBB, 10 year	0.65	(0.58)		
Nikkei 225	-1.32	(1.82)		[0.04, 0.92]
TOPIX	-0.81	(0.99)		[-0.02, 1.09]
US Dollar	-0.77	(1.80)		
Euro	0.09	(0.47)		
Panel B: Pass-Through from 7-Year JGB				
AA, 5 year	0.84**	(0.41)		[0.14, 0.59]
AA, 10 year	0.93**	(0.19)		
BBB, 5 year	0.36	(1.40)		
BBB, 10 year	2.50	(14.31)		
Nikkei 225	-0.39	(0.40)	[0.07, 0.78]	
TOPIX	-0.31	(0.30)	[0.02, 1.12]	
US Dollar	-0.09	(0.14)		
Euro	-0.02	(0.12)		
Panel C: Pass-Through from 10-Year JGB				
AA, 5 year	0.41**	(0.19)		
AA, 10 year	0.94**	(0.05)		[0.87, 0.91]
BBB, 5 year	-0.12	(0.45)		
BBB, 10 year	1.09	(1.24)		
Nikkei 225	-0.21	(0.19)		
TOPIX	-0.19	(0.18)		
US Dollar	-0.21	(0.19)		
Euro	-0.13	(0.17)		
Panel D: Pass-Through from 20-Year JGB				
AA, 5 year	0.39*	(0.23)		[0.13, 0.94]
AA, 10 year	0.84**	(0.21)	[0.65, 0.90]	[0.68, 0.88]
BBB, 5 year	0.21	(0.15)		
BBB, 10 year	-2.90	(14.96)		
Nikkei 225	-0.13	(0.08)		
TOPIX	-0.15*	(0.09)		
US Dollar	-0.21**	(0.10)		
Euro	-0.20*	(0.11)		

a. Superscripts $*, **$ denote the significance at the level of 10% and 5%, respectively. Heteroscedasticity-robust standard errors are presented in the parenthesis.

b. The 90-percent confidence interval is shown based on S and K statistics. The blank space implies that a confidence interval is $(-\infty, \infty)$.

Table 1.5: GMM Estimates of the Pass-through to Corporate Bond Yield, Stock Prices and Exchange Rates

	Estimate	Std. Error	S Conf.		K Conf.	
Panel A: Pass-Through from 5-Year JGB						
AA, 5 year	1.01**	(0.05)			[0.90,	1.19]
AA, 10 year	0.71**	(0.13)	[0.45,	0.90]	[0.52,	0.86]
BBB, 5 year	1.14**	(0.09)	[1.01,	1.36]	[-0.42,	1.29]*
BBB, 10 year	0.67**	(0.10)	[0.49,	0.97]	[0.53,	0.87]
Nikkei 225	0.29**	(0.11)	[0.11,	0.76]	[-1.56,	0.56]*
TOPIX	0.28**	(0.09)	[0.14,	0.70]	[-1.33,	0.51]*
US Dollar	-0.05	(0.07)	[-0.17,	0.04]	[-0.16,	0.04]
Euro	-0.05	(0.11)	[-0.24,	0.13]	[-0.20,	0.10]
Panel B: Pass-Through from 7-Year JGB						
AA, 5 year	1.07**	(0.12)	[1.15,	1.44]	[0.17,	1.88]*
AA, 10 year	0.45**	(0.09)	[0.24,	0.61]	[0.30,	2.00]*
BBB, 5 year	1.00**	(0.15)	[0.80,	1.44]	[-0.05,	1.30]*
BBB, 10 year	0.42**	(0.09)	[0.22,	0.64]	[0.27,	2.10]*
Nikkei 225	0.29	(0.23)	[-0.74,	2.37]	[-0.98,	-0.25]
TOPIX	0.27	(0.18)	[-0.44,	2.31]	[-0.70,	0.97]*
US Dollar	-0.13	(0.20)			[-0.23,	0.17]
Euro	-0.15	(0.26)			[-0.29,	0.24]
Panel C: Pass-Through from 10-Year JGB						
AA, 5 year	1.39*	(0.40)	[1.11,	2.06]	[0.18,	1.87]*
AA, 10 year	4.28	(7.60)			[0.75,	0.81]
BBB, 5 year	0.50**	(0.17)	[0.88,	1.72]	[0.02,	1.66]*
BBB, 10 year	0.31	(1.32)				
Nikkei 225	-0.10	(0.41)			[-1.41,	3.68]*
TOPIX	-0.05	(0.29)			[-0.66,	2.03]*
US Dollar	-0.63	(0.53)			[-0.22,	0.42]
Euro	-0.99	(1.65)				
Panel D: Pass-Through from 20-Year JGB						
AA, 5 year	0.86	(0.53)	[1.22,	3.07]	[0.12,	2.63]*
AA, 10 year	1.00*	(0.55)			[0.51,	0.63]
BBB, 5 year	0.70	(0.45)	[1.22,	3.67]	[0.06,	3.13]*
BBB, 10 year	0.14	(0.49)			[1.19,	1.44]
Nikkei 225	-0.16	(0.13)				
TOPIX	-0.13	(0.11)				
US Dollar	-0.30*	(0.16)				
Euro	-0.27**	(0.12)				

a. This table shows the pass-through to the corporate bond yields, stock prices, and exchange rates. The analysis is based on the selected announcements.

b. Superscript $\star$ indicates that the confidence interval is formed by disjoint confidence sets.

c. Other notes are the same as in Table 1.5.

Table 1.6: GMM Estimates of the Pass-through to Corporate Bond Yields, Stock Prices and Exchange Rates (Selected Announcements)

Series	Announcement	Non-Announcement	Variance Ratio
Financial Assets			
REIT index	1.65	1.52	1.18
CDS index	10.21	6.16	$2.75^{**}_{\ddagger}$
Exchange Rates: OECD Countries			
Australian Dollar	1.26	1.13	1.26
Canadian Dollar	1.23	0.94	$1.69^{**}_{\dagger}$
Korean Won	1.35	1.08	1.58^{**}
New Zealand Dollar	1.30	1.10	1.38^{*}
UK Pound	1.21	0.85	$2.05^{**}_{\dagger}$
Exchange Rates: Asian Economies			
Hong Kong Dollar	1.08	0.71	$2.32^{**}_{\ddagger}$
Singapore Dollar	0.91	0.69	$1.74^{**}_{\dagger}$
Taiwanese Dollar	1.02	0.75	$1.87^{**}_{\ddagger}$
Thai Baht	0.99	0.80	1.52^{**}

a. This table compares the standard deviation of daily changes on the announcement days and non-announcement days for additional variables: REITs, CDS, and the OECD and Asian exchange rates. Daily changes of the yield in basis points and daily percent changes of the exchange rates are used.

b. Other notes are the same as in Table 1.3.

Table 1.7: Standard Deviation and Variance Ratio of the Series on the Announcement Day and the Non-Announcement Day

	Estimate	Std. Error	S Conf.	K Conf.
Panel A: Pass-Through from 5-Year JGB				
REIT Index	-0.37	(0.27)	[-0.89, -0.40]	
CDS Index	7.57	(7.23)	[0.08, 2.26]	[2.77, 63.55]
Australian Dollar	0.17	(0.77)		
Canadian Dollar	-1.14	(2.06)		
Korean Won	-0.94	(1.53)		
New Zealand Dollar	-0.45	(1.20)		
UK Pound	0.01	(0.45)		
Hong Kong Dollar	-0.93	(2.07)		
Singapore Dollar	-0.06	(0.26)		
Taiwanese Dollar	0.02	(0.30)		
Thai Baht	-0.11	(0.43)		[-0.24, -0.03]
Panel B: Pass-Through from 7-Year JGB				
REIT Index	0.23	(0.21)	[0.17, 0.77]	[0.17, 0.74]
CDS Index	5.82	(4.29)		
Australian Dollar	-0.07	(0.13)		
Canadian Dollar	-0.18	(0.23)		
Korean Won	-0.28	(0.30)		
New Zealand Dollar	-0.13	(0.19)		
UK Pound	-0.03	(0.14)		
Hong Kong Dollar	-0.11	(0.15)		
Singapore Dollar	-0.05	(0.11)		
Taiwanese Dollar	-0.04	(0.10)		
Thai Baht	-0.04	(0.12)		

Table 1.8: GMM Estimates of the Pass-through to Additional Variables

	Estimate	Std. Error	S Conf.		K Conf.	
Panel C: Pass-Through from 10-Year JGB						
REIT Index	-0.34*	(0.19)	[-1.72,	-0.35]	[-0.71,	-0.35]
CDS Index	2.15	(1.94)				
Australian Dollar	-0.16	(0.15)				
Canadian Dollar	-0.26	(0.21)				
Korean Won	-0.19	(0.18)				
New Zealand Dollar	-0.22	(0.20)				
UK Pound	-0.16	(0.19)				
Hong Kong Dollar	-0.23	(0.19)				
Singapore Dollar	-0.16	(0.14)				
Taiwanese Dollar	-0.11	(0.13)				
Thai Baht	-0.16	(0.16)				
Panel D: Pass-Through from 20-Year JGB						
REIT Index	-0.29**	(0.12)			[-0.52,	0.27]
CDS Index	0.29	(0.41)				
Australian Dollar	-0.14*	(0.08)				
Canadian Dollar	-0.22**	(0.11)				
Korean Won	-0.19**	(0.11)				
New Zealand Dollar	-0.18**	(0.09)				
UK Pound	-0.21*	(0.12)				
Hong Kong Dollar	-0.21**	(0.10)				
Singapore Dollar	-0.16**	(0.07)				
Taiwanese Dollar	-0.15*	(0.08)				
Thai Baht	-0.16**	(0.08)				

a. This table shows the pass-through to additional variables: REITs, CDS, and the OECD and Asian exchange rates.

b. Other notes are the same as in Table 1.5.

Table 1.8: GMM Estimates of the Pass-through to Additional Variables (continued)

	Estimate	Std. Error	S Conf.		K Conf.	
Panel A: Pass-Through from 5-Year JGB						
REIT Index	0.32*	(0.10)	[0.17,	0.53]	[-1.57,	0.48]⋆
CDS Index	-0.93	(0.58)	[-2.23,	0.21]	[-1.86,	-0.05]
Australian Dollar	-0.03	(0.07)	[-0.31,	0.19]	[-0.18,	0.10]
Canadian Dollar	-0.03	(0.08)	[-0.22,	0.18]	[0.10,	1.18]
Korean Won	0.02	(0.06)	[-0.14,	0.27]	[-0.08,	0.16]
New Zealand Dollar	0.01	(0.08)	[-0.18,	0.35]	[-0.12,	0.18]
UK Pound	-0.07	(0.11)	[-0.42,	0.05]	[-0.27,	0.03]
Hong Kong Dollar	-0.03	(0.07)	[-0.13,	0.08]	[-0.27,	0.03]
Singapore Dollar	-0.06	(0.08)	[-0.20,	0.04]	[-0.13,	0.08]
Taiwanese Dollar	-0.05	(0.07)	[-0.19,	0.04]	[-0.17,	0.03]
Thai Baht	-0.05	(0.07)	[-0.26,	0.05]	[-0.16,	0.03]
Panel B: Pass-Through from 7-Year JGB						
REIT Index	0.32**	(0.09)	[0.21,	0.54]	[-0.70,	0.49]⋆
CDS Index	-1.40	(1.02)	[-4.02,	0.04]	[-2.98,	-0.20]
Australian Dollar	0.07	(0.20)				
Canadian Dollar	-0.00	(0.14)				
Korean Won	0.17	(0.18)				
New Zealand Dollar	0.06	(0.19)				
UK Pound	-0.09	(0.21)				
Hong Kong Dollar	-0.12	(0.18)			[-0.20,	0.21]
Singapore Dollar	-0.12	(0.18)			[-0.24,	0.18]
Taiwanese Dollar	-0.10	(0.18)			[-0.24,	0.21]
Thai Baht	-0.09	(0.18)				

Table 1.9: GMM Estimates of the Pass-through to Additional Variables (Selected Announcements)

	Estimate	Std. Error	S Conf.		K Conf.	
Panel C: Pass-Through from 10-Year JGB						
REIT Index	-0.07	(0.21)	[0.15,	1.57]	[-0.51,	1.27]*
CDS Index	1.07	(1.22)				
Australian Dollar	-0.22	(0.45)				
Canadian Dollar	-0.29	(0.27)				
Korean Won	0.08	(0.45)				
New Zealand Dollar	-0.25	(0.36)				
UK Pound	-0.47	(0.36)				
Hong Kong Dollar	-0.54	(0.41)			[-0.19,	0.34]
Singapore Dollar	-0.63	(0.66)				
Taiwanese Dollar	-0.63	(0.69)				
Thai Baht	-0.45	(0.34)				
Panel D: Pass-Through from 20-Year JGB						
REIT Index	-0.20**	(0.09)	[0.00,	1.49]	[-0.38,	1.24]*
Australian Dollar	-0.14	(0.09)				
CDS Index	0.32**	(0.09)				
Canadian Dollar	-0.19**	(0.07)				
Korean Won	-0.09	(0.12)			[-0.25,	-0.05]
New Zealand Dollar	-0.18**	(0.08)				
UK Pound	-0.24**	(0.07)				
Hong Kong Dollar	-0.30**	(0.15)				
Singapore Dollar	-0.25**	(0.12)				
Taiwanese Dollar	-0.25**	(0.13)				
Thai Baht	-0.25**	(0.10)				

a. This table shows the pass-through to additional variables: REIT, CDS, and the OECD and Asian exchange rates. The analysis is based on the selected announcements.
b. Other notes are the same as in Tables 1.5 and 1.6.

Table 1.9: GMM Estimates of the Pass-through to Additional Variables (Selected Announcements) (continued)

	2001-2006 (QE)	2006-2010	2010-2013 (CME)
Panel A: Pass-Through from 5-Year JGB			
AA, 5 year	1.02**	1.10**	1.13
AA, 10 year	1.07**	0.95**	3.64
BBB, 5 year	1.04**	1.13**	1.58
BBB, 10 year	0.79**	1.08**	3.09
Nikkei 225	-0.66	-0.34	0.71
TOPIX	-0.53	-0.28	0.64
US Dollar	-0.04	-0.06	-0.04
Euro	0.00	-0.07	0.05
Panel B: Pass-Through from 7-Year JGB			
AA, 5 year	0.76**	0.33**	0.55**
AA, 10 year	0.96**	0.55	1.28**
BBB, 5 year	1.01**	0.44*	0.87**
BBB, 10 year	0.49**	1.21	1.86**
Nikkei 225	-0.07	-0.94	0.31
TOPIX	-0.06	-0.62	0.32
US Dollar	-0.02**	-0.08	0.01
Euro	0.05	-0.24	0.06
Panel C: Pass-Through from 10-Year JGB			
AA, 5 year	0.73**	1.45	0.79**
AA, 10 year	0.98**	0.99**	1.80**
BBB, 5 year	1.09**	1.68	1.13**
BBB, 10 year	0.60*	0.84	1.98**
Nikkei 225	-0.05	1.61	0.48
TOPIX	-0.04	0.82	0.46
US Dollar	-0.03**	-0.21	-0.03
Euro	0.05	-0.33	0.12
Panel D: Pass-Through from 20-Year JGB			
AA, 5 year	0.55**	0.98**	0.36**
AA, 10 year	0.97**	0.94**	0.75**
BBB, 5 year	1.19**	1.07**	0.52**
BBB, 10 year	0.44**	1.13**	1.08**
Nikkei 225	-0.06	0.79	0.11
TOPIX	-0.06	0.39	0.09
US Dollar	-0.02	-0.18	0.04
Euro	0.05	-0.20	-0.02
Number of Announcements	13	13	9

a. This table shows the pass-through to the corporate bond yields, stock prices, and exchange rates in the subsamples: 2001-2006 (the QE program), 2006-2010, and 2010-2013 (the CME program).

b. Other notes are the same as in Table 1.5.

Table 1.10: GMM Estimates of the Pass-through in Subperiods

	Estimate	Std. Error	S Conf.		K Conf.	
Panel A: Pass-Through to Main Variables						
AA, 5 year	0.42**	(0.16)	[0.32,	0.50]	[0.32,	0.45]
AA, 10 year	0.58**	(0.17)				
BBB, 5 year	-0.09	(0.50)				
BBB, 10 year	0.75	(0.85)				
Nikkei 225	-0.15	(0.23)				
TOPIX	-0.12	(0.18)				
US Dollar	-0.04	(0.07)				
Euro	0.00	(0.05)				
Panel B: Pass-Through to Additional Variables						
REIT Index	0.13	(0.19)	[0.03,	0.96]	[0.05,	0.58]
CDS Index	6.04	(8.00)				
Australian Dollar	-0.04	(0.07)				
Canadian Dollar	-0.10	(0.15)				
Korean Won	-0.15	(0.22)			[0.00,	0.49]
New Zealand Dollar	-0.06	(0.11)				
UK Pound	-0.01	(0.06)				
Hong Kong Dollar	-0.05	(0.09)				
Singapore Dollar	-0.03	(0.05)				
Taiwanese Dollar	-0.01	(0.04)				
Thai Baht	-0.03	(0.05)				

a. This table presents the pass-through based on the principal component of JGB yields with a maturity of 5,7,10 and 20 years.

b. Other notes are the same as in Table 1.5.

Table 1.11: GMM Estimates of the Pass-through based on the Principal Component of JGB Yields

Chapter 2

Evaluating the Efficiency of the FOMC's New Economic Projections

2.1 Introduction

Since October 2007, the Federal Open Market Committee (FOMC) has been publishing the "Summary of Economic Projections" after the meetings. These economic projections are numerical projections of four macroeconomic series submitted by individual FOMC policymakers. With a few years of these new projections in hand, researchers can now begin to assess their efficiency.

The assessment of FOMC's new economic projections is important for two reasons. First, given the findings that subjective forecasts are often more accurate than forecasts using reduced-form models,[1] analyzing the efficiency of another subjective forecast is a subject of interest. In particular, these projections are based on profound knowledge and judgment about economics since the FOMC has made considerable efforts to make them more accurate and consistent with economic narratives. Second, the assessment of the FOMC's new

[1]For details, see Ang, Bekaert, and Wei (2007) and Faust and Wright (2009).

projections is important in practice because monetary policy decisions are now explicitly tied to these projections and the public is keenly aware of them.[2]

In this paper, I evaluate the efficiency of FOMC's new projections by testing if their revisions are unpredictable. Even though the efficiency could be evaluated by testing the unpredictability of forecast errors, the power of the tests would be very low due to the short sample in hand. Therefore, I focus on the unpredictability of forecast revisions. I propose different tests based on the time-series property of forecast revisions (bias, autocorrelation, and Wald statistic) and signs of forecast revisions (positive revisions and consecutive revisions). In addition, I propose the joint tests across different target years and series to improve the power of the tests.

One limitation of this analysis is that the period in which these projections have been made (2007—2014) is very short, and contains an extremely turbulent period for the US economy. Forecasting a macroeconomic series is difficult even during the normal times, and evaluating the efficiency by looking at this particularly unusual period may not be appropriate. However, an evaluation in this early stage can still be a useful benchmark, considering close attention paid to the new projections. To evaluate the size and power of the tests in the small sample, I provide a Monte Carlo exercise. The simulation results show that the size is generally close to the nominal size (though the joint tests tend to be slightly oversized), and some tests are powerful even with a small sample, against a reasonable set of simulations where the forecasts are not efficient.

The results show a stark contrast between the forecast efficiency of real

[2]For example, in December 2012, Bernanke (2012a) explains that the FOMC's decision to use their unemployment projections to give guidance on how long they will keep the federal fund rate low is to "make FOMC's intention to maintain accommodation more explicit."

economic projections and inflation projections. While the efficiency is accepted for inflation in almost all years, it is often rejected for real economic variables, notably for the unemployment rate. For unemployment, the rejections between 2009 and 2011 are so strong that they lead to the rejections in the joint tests.[3] The joint efficiency of the entire projection is accepted in most cases.

In order to compare the results, I apply the same tests to the Survey of Professional Forecasters (SPF) forecast. Similar to the FOMC's projections, the efficiency of the SPF's inflation forecast is accepted in most cases. On the other hand, the SPF's unemployment forecast is not as inefficient as FOMC's unemployment projections. This comparison highlights that the revisions of FOMC's unemployment projections have a much stronger autocorrelation, which may reflect information rigidity in FOMC's unemployment projections.

This strong rigidity in FOMC's unemployment projections is puzzling for two reasons. First, it is not consistent with Okun's law, which draws an negative association between unemployment and GDP growth. Second, the FOMC's projections should in principle be at least as efficient as the SPF forecast, since they have an access to the Greenbook forecast—a forecast that is prepared by the staff of the Federal Reserve and generally more accurate than the SPF forecast. In order to facilitate the accurate interpretation of these results, I discuss the following explanations: (1) slower updating of FOMC's beliefs about unemployment, (2) FOMC's conservatism about their unemployment projections, and (3) different predictability of GDP growth and the unemployment rate.

Lastly, I provide an extension to the main results. I analyze the relationship

[3]For example, projections for the fourth quarter of 2009 are all revised upward through 2007 to 2009, as described in Figure 2.1, which is highly unlikely under the null hypothesis of unpredictable forecast revisions.

between the revisions of the FOMC's GDP growth and unemployment projections to assess if they follow Okun's law. The regression analysis shows that the FOMC's projections are consistent with Okun's law only at the shorter horizons.

The remainder of the paper is organized as follows: Section 2.2 explains the projections I use, and Section 2.3 describes the method of forecast evaluations and provides a Monte Carlo exercise. Section 2.4 contains the main results and Section 2.5 provides a regression analysis focusing on Okun's law as an extension. Section 2.6 concludes.

2.2 Data

The FOMC's economic projections are numerical projections of four macroeconomic series, real GDP growth, the unemployment rate, PCE inflation and Core PCE inflation, over next two to three years. Each FOMC policymaker submits his/her own projections at the FOMC meeting, and the range and the central tendency of projections are published after the meeting. The central tendency is the range of projections from which the three highest and lowest projections are excluded. In this paper, I focus on the midpoints of the central tendency and range to analyze the revisions of these projections.[4]

The FOMC has introduced these projections as a part of the Fed's enhanced communication strategy.[5] In many respects, they are considerably more detailed and informative than the old projections that had been released twice a year with the monetary policy report to US Congress. First, these new projections

[4]The results based on the upper end or the lower end of the projections are generally similar as the results using the midpoint of the projections. Gavin and Pande (2008) and Fischer, García-Bárzana, Tillmann, and Winker (2014) discuss the implication and caveats of taking the midpoints of intervals of FOMC's projections.

[5]For details, see Bernanke (2007) and Mishkin (2007).

have been published more frequently than the old projections, four times a year, in March, June, September, and December.[6] Because FOMC policymakers have a chance to revise their projections right after the meeting, these projections can be regarded as the forecasts conditional on the information set of the meeting day.

Second, the horizons of new projections have been extended to three years; the old projections had a two-year horizon. The new projections aim to forecast the level of unemployment rate in the fourth quarter, and the rate of changes of real GDP and prices in the fourth quarter from a year earlier.

There are at most fourteen projections, and thus thirteen consecutive revisions, for an individual target year. Since these projections are newly introduced, the number of revisions for some target year is limited. In my dataset, I have eight revisions for 2009, twelve revisions for 2010 and 2011, thirteen revisions for 2012 and 2013, and eleven revisions for 2014. The projections for 2008 and 2015 are dropped because the sample is too small.

Even though more detailed distribution of FOMC's projections is published three weeks after the meeting, it is anonymous and I cannot match individual policymaker's responses across different series and periods. Accordingly, I primarily focus on the central tendency and range of the projections.

2.3 Method

This section describes the tests I use to evaluate the efficiency of FOMC's projections. Based on an implication of forecast efficiency that the revisions are

[6] The projections used to be released typically in January, April, June, and November. But the FOMC has changed the timing since 2012, by releasing five projections in 2012.

unpredictable, I propose several tests focusing on time-series property or signs of forecast revisions. In addition, I propose the joint tests using the average across target years and series as the test statistics. Then, I describe the inference based on the bootstrap. Lastly, I provide a Monte Carlo exercise assessing the size and power of the tests to show that small sample issues are not too pronounced.

2.3.1 Testable Implication of Forecast Efficiency

The idea of a forecast efficiency test using revisions dates back to Nordhaus (1987). Essentially, it is based on the implication that forecast revisions are unpredictable.[7] To formalize the idea, suppose that y_t is a variable of interest, and denote $\hat{y}_{t+h|t}$ as a forecast for period $t+h$, based on the set of variables observed in period t, $\mathbf{X_t}$. Then define the forecast revision for period $t+h$ between t and $t+j$, for any j such that $0 < j < h$, as $r_{t+h|t,t+j} \equiv \hat{y}_{t+h|t+j} - \hat{y}_{t+h|t}$.

It is well known that the optimal forecast is the conditional expectation of the series under a symmetric loss function. Therefore, the realized value at period $t+h$ is the sum of the conditional expectation, $E[y_{t+h}|\mathbf{X}_t]$, and its uncorrelated forecast error, $e_{t+h|t}$:

$$y_{t+h} = E[y_{t+h}|\mathbf{X}_t] + e_{t+h|t}. \tag{2.1}$$

Then, the revision between t and $t+j$ is described as the difference between

[7] Isiklar, Lahiri, and Loungani (2006) also evaluate the efficiency of Consensus Economics forecast by testing the unpredictability of revisions. For more recent application, see Loungani, Stekler, and Tamirisa (2013) and Sheng (forthcoming).

forecast errors in period t and $t+j$:

$$r_{t+h|t,t+j} = E[y_{t+h}|\mathbf{X}_{t+j}] - E[y_{t+h}|\mathbf{X}_t],$$

$$= e_{t+h|t} - e_{t+h|t+j}. \tag{2.2}$$

Since $e_{t+h|t}$ is also uncorrelated to $\mathbf{X}_{t+j}$, $r_{t+h|t,t+j}$ is uncorrelated to $\mathbf{X}_{t+j}$. As a result, revisions of the efficient forecasts are uncorrelated to any observable variables. By setting the maximum forecast horizon as H, the sequence of consecutive forecast revisions are described as $\{r_{t+H|t+h-1,t+h}\}_{h=0}^{H-1}$. In this paper, I primarily focus on this sequence of forecast revisions.[8]

2.3.2 Test for Individual Year and Series

In order to test the efficiency of FOMC's projections for an individual target year and series, I propose the methods focusing on two different properties of forecast revisions: *time-series properties* and *signs* of forecast revisions.

2.3.2.1 Tests Using Time-Series Properties of Revisions

For the tests using time-series properties of forecast revisions, I use three summary statistics: (1) Bias, (2) First-order autocorrelation, and (3) Wald statistic of the first-order autoregression. To define these statistics formally, consider a first-order autoregression of forecast revisions:

$$r_{t+H|t+h-1,t+h} = \alpha_{t+H} + \beta_{t+H} \cdot r_{t+H|t+h-2,t+h-1} + \varepsilon_{t+h}. \tag{2.3}$$

The forecast efficiency implies both the intercept, α_{t+H}, and the coefficient, β_{t+H}, are zero. The first test statistic, the bias of forecast revisions, tests if

[8] As pointed out by Patton and Timmermann (2012), incorporating forecast revisions will make the forecast evaluation significantly more powerful, which could be even used to improve the accuracy of forecasts as discussed in Arai (2014).

$\alpha_{t+H} = 0$. The sample bias is computed as the average of forecast revisions:

$$\bar{r}_{t+H} = \frac{1}{H} \sum_{h=0}^{H-1} r_{t+H|t+h-1,t+h}. \tag{2.4}$$

The second test statistic, the first-order autocorrelation of forecast revisions, tests if $\beta_{t+H} = 0$. The sample autocorrelation is computed as the ratio of autocovariance to its variance:

$$\hat{\rho}^1_{t+H} = \frac{\hat{\gamma}^1_{t+H}}{\hat{\gamma}^0_{t+H}}, \tag{2.5}$$

where

$$\hat{\gamma}^j_{t+H} = \frac{1}{H} \sum_{h=0}^{H-j-1} (r_{t+H|t+h-1,t+h} - \bar{r}_{t+H})(r_{t+H|t+h+j-1,t+h+j} - \bar{r}_{t+H}) \ \text{ for } \ j = 0, 1.$$

I use the sample mean $\bar{r}_{t+H}$ to measure the deviations of lagged series, and the total number of revisions H to normalize.

The third test statistic, the Wald statistic, jointly tests if $\theta_{t+H} \equiv [\alpha_{t+H}, \beta_{t+H}]'$ is a zero vector. The sample Wald statistic is computed as follows:

$$\hat{W}_{t+H} = H \ \hat{\theta}'_{t+H}[(Avar(\hat{\theta}_{t+H}))]^{-1}\hat{\theta}_{t+H}, \tag{2.6}$$

where $Avar(\hat{\theta}_{t+H})$ is the asymptotic variance-covariance matrix of $\hat{\theta}_{t+H}$. I estimate the asymptotic variance without any autocorrelation correction because the forecast efficiency implies that revisions are serially uncorrelated.

2.3.2.2 Tests Using Signs of Revisions

For the tests using signs of forecast revisions, I use two summary statistics: (1) Ratio of positive forecast revisions and (2) Ratio of the cases in which the consecutive forecast revisions have the same sign. The advantage of focusing on signs is that it gives the exact distribution of test statistics, which enables us to do the exact test.

The first test statistic summarizes how often the forecast revision is positive. Since the sign of revisions can be regarded as an outcome of the Bernoulli trial under the forecast efficiency, the number of positive revisions should follow *Binomial* $(H, 0.5)$. I divide it by H to normalize as the ratio.

To define the test statistic, first define the indicator variable $i^P_{t+H|t+h}$ for a target period of $t + H$:

$$i^P_{t+H|t+h} = \begin{cases} 1 & \text{if} \quad r_{t+H|t+h-1,t+h} > 0 \ \text{ for } \ h = 0, \ldots, H-1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.7}$$

Then, the test statistic is defined as the ratio of the sum of these indicator variables to the total number of forecast revisions:

$$b^P_{t+H} = \frac{1}{H} \sum_{h=0}^{H-1} i^P_{t+H|t+h}. \tag{2.8}$$

Similarly, I can define the second test statistic, which summarizes how often the consecutive forecast revisions have the same sign, as being either positive or negative. Since such event can also be regarded as an outcome of Bernoulli trial under the forecast efficiency, the number of such cases should follow *Binomial* $(H - 1, 0.5)$. I divide it by $H - 1$ to normalize as the ratio.

Let $i^C_{t+H|t+h}$ be the indicator variable for a target period of $t + H$:

$$i^C_{t+H|t+h} = \begin{cases} 1 & \text{if} \quad r_{t+H|t+h-1,t+h} \cdot r_{t+H|t+h,t+h+1} > 0 \ \text{ for } \ h = 0, \ldots, H-2, \\ 0 & \text{otherwise.} \end{cases} \tag{2.9}$$

Then, the test statistic is defined as the ratio of the sum of these indicator variables to the total number of consecutive forecast revisions:

$$b^C_{t+H} = \frac{1}{H-1} \sum_{h=0}^{H-2} i^C_{t+H|t+h}. \tag{2.10}$$

When a value of the forecast revision is zero, I compute these test statistics in two steps. First, I randomly assign a sign with the probability of 0.5 to compute these statistics. Second, I repeat this random assignment many times (100 times in this paper) to treat the mean as the test statistic.

One concern about focusing on signs is that it may over-simplify the forecast revisions, and thus the tests may not have enough power. However, as presented in a Monte Carlo exercise and the empirical results, these tests reject the null as many cases as the tests based on time-series properties, which suggests that the loss of the power associated with this simplification is not detrimental.[9]

2.3.3 Joint Test Across Years and Series

One limitation of the efficiency tests for an individual year and series is that they may not have enough power due to the short sample. In order to make tests more powerful, I compute the joint test statistics across different years and series by averaging individual test statistics.

First, I compute the average of individual test statistics across different target years. Define the vector of individual test statistics for the target year t as $\mathbf{x}_t \equiv \{\bar{r}_t, \hat{\rho}_t^1, \hat{W}_t, b_t^P, b_t^C\}$ for some series. Then the vector of joint test statistics from year T_s to year T_e is defined as the average of individual test statistics:

$$\bar{\mathbf{x}} \equiv \frac{1}{T_e - T_s + 1} \sum_{t=T_s}^{T_e} \mathbf{x}_t. \tag{2.11}$$

Second, I compute the average of these statistics across all series as the test statistic for the entire projections. The formal expression is abbreviated to conserve space.

[9] Campbell and Ghysels (1995, 1997) also apply the similar tests to budget forecasts in the US and Canada, and claim that these tests have good finite sample power.

2.3.4 Inference

I conduct the exact tests for the individual tests using signs, and approximate tests based on the bootstrap for other individual tests and joint tests. I report one-sided p-values for the Wald statistics and their averages, and two-sided p-values for other test statistics.

The bootstrap p-values are computed based the null hypothesis that the FOMC's projections are efficient, and therefore their revisions are serially uncorrelated. However, to keep the correlation among the forecast revisions at different horizons, I construct the artificial projections by resampling a block of multiple forecast revisions made at the same period. In addition, I apply the wild bootstrap to the FOMC's forecast revisions, in which I flip the sign of the resampled revisions with the probability of 0.5. This is because the distribution of forecast revisions are symmetric under the null hypothesis.

After constructing artificial projections, I apply the same tests to obtain the bootstrap test statistics. By repeating this procedure arbitrarily many times, I can form the distribution of bootstrap test statistics and report the p-value based on the percentile of the sample test statistics.

2.3.5 Monte Carlo Exercise

I conduct a Monte Carlo exercise to assess the size and power of the tests. The simulation results show that the size is generally close to the nominal size and some tests are powerful even with a small sample.

By constructing artificial forecast revisions using a reduced-form VAR, I first check if the actual probability of rejections is close to the nominal size.

Then, I construct three types of inefficient forecasts to assess the power: (1) Forecasts with the independent noise, (2) Forecasts with the persistent noise across multiple horizons, and (3) Forecasts with the sluggish adjustments.[10] I use the data from 1984 to 2012 to calibrate the series, which includes both the Great Moderation and the period after the Great Recession, but excludes the period before the Great Moderation. The specific steps are described in detail in the Appendix B.

The simulation results with the nominal size of 10% are presented in Tables 2.1 to 2.4. The size is generally close to the nominal size in most individual tests. However, the consecutive sign test tends to be slightly oversized, and the joint test across different years and series tends to be oversized, with the actual size up to 23.6%.

For the power of the tests, the results show that the autocorrelation, Wald, and consecutive sign tests are much more powerful than the other two tests, and their higher power is robust across different simulations.[11] In particular, the autocorrelation and consecutive sign tests have an extremely high power close to 1, in all joint tests. These results show that some tests used in this paper are considerably powerful even with a small sample, against a set of reasonable simulations where the forecasts are not efficient.

[10] The variance of the noise is set to be unity in both the first and second simulations. However, the results are generally similar when the sample variance of each series is used.

[11] However, the poor power of the bias and positive sign tests is primarily due to the design of the simulations, in which all inefficient forecasts have the same mean as the efficient forecasts.

2.4 Results

In this section, I present the efficiency evaluation of FOMC's projections, showing that the efficiency is rejected often for real economic variables, especially for the unemployment rate, while it is accepted for inflation. Then I compare the results with the SPF forecasts to highlight that the revisions of FOMCs unemployment projections have a much stronger autocorrelation, which may suggest information rigidity of the projections. Then I discuss that slow updating, conservatism, or different predictability could help explain such rigidity.

2.4.1 FOMC's Economic Projections

The results using the midpoints of the central tendency and the range are presented in Tables 2.5 and 2.6, respectively. The results show that there is a stark contrast between the forecast efficiency of real economic projections and inflation projections. While the efficiency is accepted for inflation in almost all target years, it is rejected in many cases for real economic variables, notably for the unemployment rate. In particular, for the unemployment rate, the efficiency is rejected in most individual tests for the target years of 2009–2011. Furthermore, the joint efficiency is rejected in most joint tests across different target years because the rejections between 2009 and 2011 are so strong. Unlike the case of unemployment projections, the efficiency of output growth projections is accepted in most cases. The joint efficiency of the entire projections is accepted in most cases, and both the central tendency and the range of the projections provide similar results.

2.4.2 SPF Forecasts

The results of the SPF forecasts using the mean and the median are presented in Tables 2.7 and 2.8, respectively. Similar to the FOMC's projections, the efficiency of the SPF's real GDP growth and unemployment forecasts is rejected more often than inflation forecasts. However, the efficiency of the SPF's unemployment forecast is rejected only in 2009, which is not as strong as the rejections of FOMC's projections.

The comparison between the FOMC's projections and the SPF forecast shows that the revisions of FOMC's unemployment projections have much stronger autocorrelations. For example, the autocorrelations of the revisions for 2010 and 2011 are 0.00 and -0.03 for the SPF forecast (mean) but 0.53 and 0.48 for FOMC's projections (central tendency), respectively.

Such significant autocorrelations in the revisions may reflect information rigidity in FOMC's unemployment projections. For example, Coibion and Gorodnichenko (2010) observe a substantial degree of correlation among the SPF's forecast revisions, and consider it as evidence of information rigidity. Coibion and Gorodnichenko (2012) also show that a broad range of survey forecasts substantially deviate from the null hypothesis of full information, and show that their finding is consistent with the predictions of macroeconomic models with information rigidity. However, it should be noted that the FOMC's projections and the SPF forecast forecast differ slightly with respect to their targets and horizons,[12] and these differences may lead to their divergent evaluations of

[12]The SPF forecasts are different from the FOMC's projections in two ways. First, the availability of longer-horizon forecasts are limited; Three-year ahead SPF forecasts starts from 2009Q2 for GDP growth and unemployment, and two-year ahead SPF forecasts start from 2007Q1 for inflation. Second, the SPF forecasts have different targets for GDP growth

forecast efficiency.

2.4.3 Discussion

The strong rigidity of FOMC's unemployment forecasts is puzzling for two reasons. First, it is inconsistent with Okun's law, which draws an negative association between unemployment and GDP growth. In other words, if forecasters follow Okun's law, evaluations of GDP growth forecasts and unemployment forecasts should be similar.[13] Second, the FOMC's projections should be in principle at least as efficient as the SPF forecast, because the FOMC has an access to the Greenbook forecast, which is generally more accurate than the SPF forecast. To facilitate the accurate interpretation of these results, I list a number of possible explanations.

First, FOMC policymakers may have gradually learned of the effect of the Great Recession on the unemployment rate over the course of multiple years. As a result, the updating of their beliefs about unemployment happens at a rate slower than that about GDP growth. Considering that FOMC policymakers have made substantial revisions to their unemployment projections at time-horizons exceeding one year, slower updating could be a likely cause of inefficiency.

Second, FOMC policymakers may be conservative about their projections

and unemployment; the SPF forecasts aim to forecast the *annual average* rate of GDP growth and level of unemployment.

[13] For the analysis of Okun's law during the Great Recession, see Elsby, Hobijn, and Sahin (2010) and Daly and Hobijn (2010). Some FOMC policymakers, such as Yellen (2010) and Bernanke (2012b), observe that both the jump in the unemployment rate in 2009 and its decline in 2011 were not anticipated by Okun's law. Ball, Leigh, and Loungani (2012) examine cross-country data and argue that Okun's law did not substantially change during the Great Recession. Daly, Fernald, Òscar Jordà, and Nechio (2014) claim that the deviation is not substantial considering the recent data revisions.

for various reasons. For example, they may be concerned about the signalling value of their projections. In other words, FOMC policymakers may be cautious about changing their projections, because such changes would convey a message about future economic conditions. As a result, this concern may lead to smoothing in their projections. In addition, FOMC policymakers may focus on worst-case scenarios in their projections.[14] More specifically, even if the FOMC policymakers were to correctly recognize developments within the labor market, they may be conservative in updating their beliefs, because their recognition is subject to misjudgment in real time. Such considerations may also lead to the inefficiency in their projections. Other factors—such as strategic behavior among committee members or concern for their reputations as forecasters—could also be causes of FOMC's conservatism. Using a dataset of FOMC's old projections presented by Romer (2010), which includes the individual forecasts of each policymaker, Nakazono (2013), Rülke and Tillmann (2011), and Tillmann (2011) each point out that the forecasting behavior of FOMC members varies with their status within the committee (i.e., governors vs. voting members vs. nonvoting members). For example, they find that governors tend to have views close to the consensus whereas non-governors tend to have extreme views, which suggests that FOMC members strategically use their forecasts to influence FOMC's decisions. Jain (2013) suggests that forecasters' reluctance to make substantial revisions is sufficiently strong to lead to substantial forecast

[14] Responding to Romer and Romer's (2008) criticism of the FOMC in terms of its inferior forecasting performance relative to the Greenbook forecast, Ellison and Sargent (2012) provide a defense by allowing the FOMC to doubt the economic model that underpins the Greenbook forecast. They claim that it is inappropriate to evaluate FOMC's forecasting performance by using the same metric, because the FOMC's and Greenbook's forecasting objectives are different.

smoothing.

Finally, differences in predictability between output growth and the unemployment rate may lead us to reject the efficiency of unemployment projections more frequently. As Tulip (2009) points out, output growth has essentially become unpredictable after the Great Moderation. Consistent with his argument, the output growth and unemployment rates simulated in the Monte Carlo exercise in Section 2.3.5 also exhibit substantially different predictability. When projecting the artificial realized series on the conditional expectation, average R-square is 0.202 for the output growth whereas 0.960 for the unemployment rate. Since the forecast evaluation proposed in this paper tests the unpredictability of forecast revisions, it could be easier for us to detect inefficiency in more predictable series, namely the unemployment rate.

2.5 Extension

To shed some light on the reason why the rejections of efficiency are stronger for the unemployment rate than GDP growth, I provide a regression analysis to assess if the FOMC's projections numerically follow Okun's law. Since Okun's law conventionally associates 1% faster GDP growth from the potential with 0.5% decline in the unemployment rate, the slope coefficient of the unemployment revisions on output growth revisions should be around -0.5. In addition to the analysis of revisions at individual horizons, I provide the analysis of cumulative revisions, which add up he revisions at all horizons, to see if these projections follow Okun's law in a medium term.

The confidence interval is computed by the block bootstrap, in which the artificial sample of a year is constructed by resampling from the sample of the same year. For example, an artificial sample of 2009 is made by resampling only from the sample of 2009. This is because the news at the current period could influence the forecasts both at shorter and longer horizons, and therefore its effect would likely to persist more than a year. As a result, I need to keep the correlation between the current revision and the revisions made more than a year before.

The results presented in Table 2.9 show that the FOMC's projections are consistent with Okun's law at shorter horizons, but inconsistent at the longer horizon. For nowcasts and 1-year ahead projections, correlations are negative: -0.41 and -0.81 for the central tendency and -0.41 and -0.72 for the range, respectively. On the other hand, the correlation in 2-year ahead projections is positive but the confidence interval is extremely wide. This is because FOMC participants substantially revise their 2-year ahead unemployment projections without changing the output growth projections. As a result, R^2 of the regression becomes very small. Furthermore, this large variation in unemployment revisions at 2-year horizon makes the slope coefficient of cumulative revisions strongly negative, -0.91 for the central tendency and -0.89 for the range. Figures 2.2 and 2.3 show the plots of the FOMC's GDP growth revisions and unemployment revisions based on the central tendency and the range, respectively, highlighting the difference between the projections at shorter and longer horizons.

2.6 Conclusion

In this paper, I evaluate the efficiency of FOMC's new economic projections by testing if their forecast revisions are unpredictable. These projections are released from 2007, and play an increasingly important role in formulation of U.S. monetary policy. Therefore, evaluating the quality of these projections is a matter of great importance for macroeconomists.

By applying several statistical tests focusing on the unpredictability of forecast revisions, I find that the efficiency of FOMC's projections is accepted for inflation in almost all target years, but often rejected for real economic variables, notably for the unemployment rate. Furthermore, the comparison with the SPF forecast shows that the inefficiency of FOMC's unemployment projections is due to the strong autocorrelation in revisions, which may reflect information rigidity in their unemployment projections.

I discuss that such strong rigidity may be related to three factors: slower updating of FOMC's beliefs about unemployment, FOMC's conservatism about their unemployment projections, and different predictability of GDP growth and the unemployment rate. Further research on disentangling these channels, especially using disaggregated data if it becomes public in the future, would yield a much better understanding of FOMC's decision making and their conduct of U.S. monetary policy.

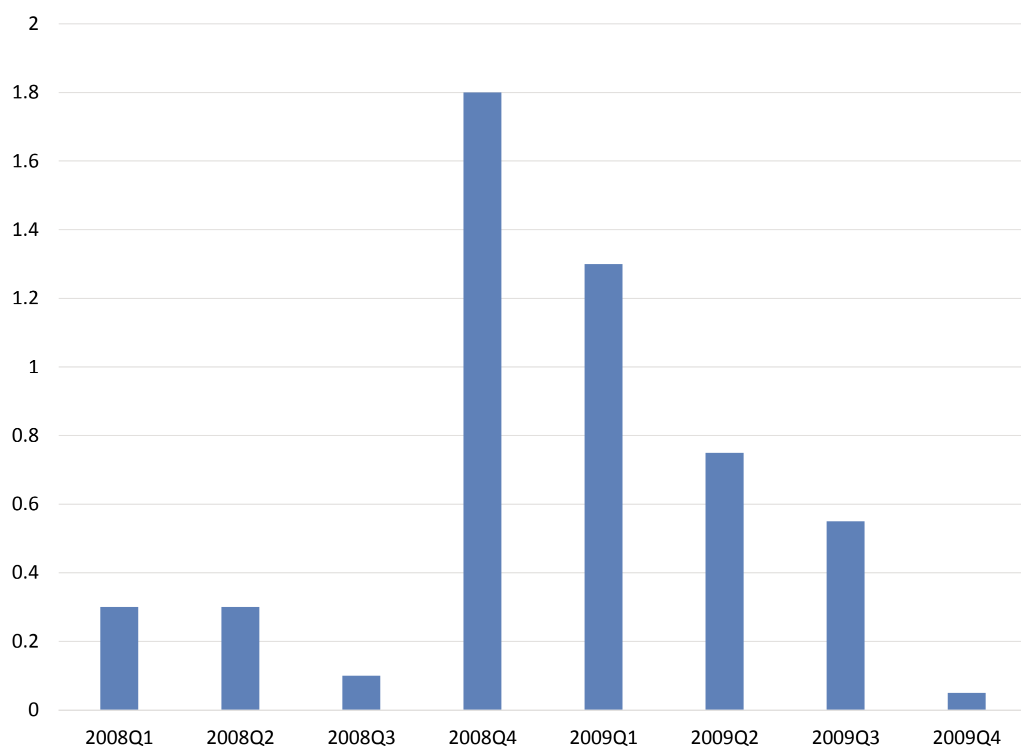

Figure 2.1: Revisions to FOMC's Unemployment Projections (Targeting 2009Q4)

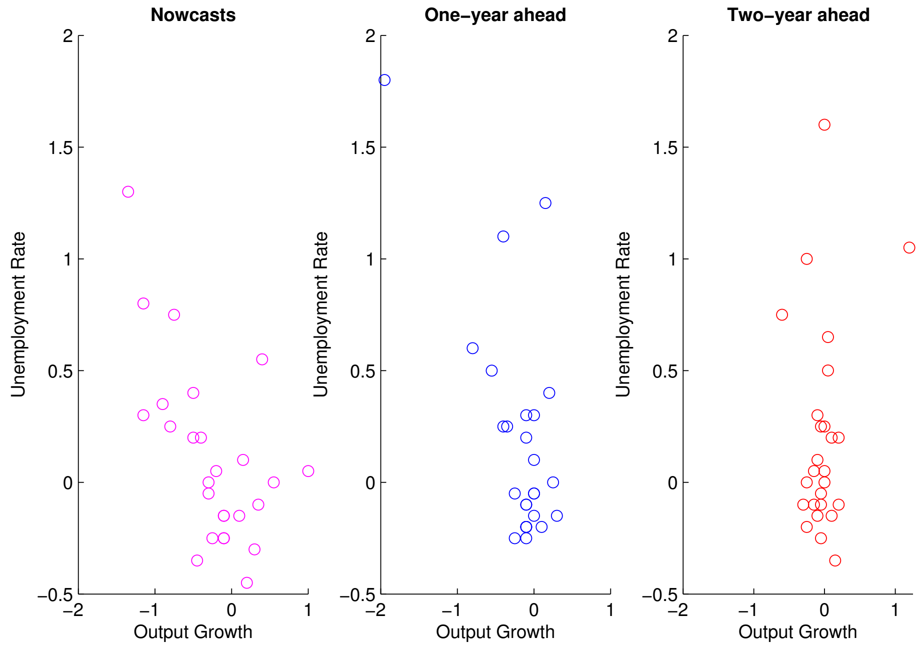

Figure 2.2: Revisions to FOMC's GDP Growth and Unemployment Projections (Central Tendency)

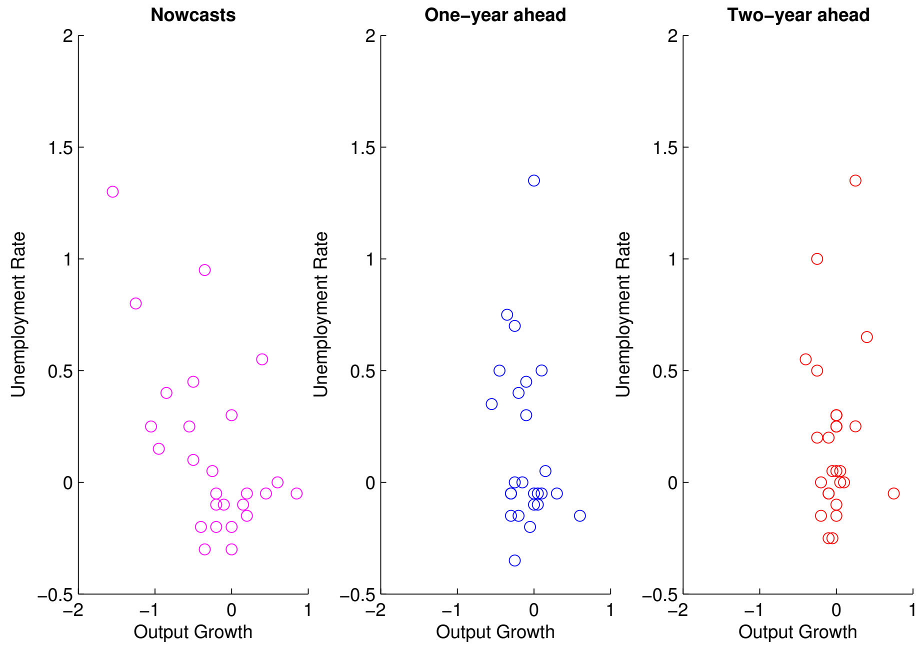

Figure 2.3: Revisions to FOMC's GDP Growth and Unemployment Projections (Range)

Projected Year	Bias	Autocorrelation	Wald	Signs, Positive	Signs, Consecutive
Panel A: Size of the Test (Individual Year)					
GDP growth	0.099	0.126	0.117	0.089	0.150
Unemployment	0.096	0.105	0.090	0.087	0.144
PCE	0.105	0.113	0.096	0.101	0.148
Core PCE	0.096	0.097	0.087	0.102	0.125
Panel B: Size of the Tests (Average of Five Years)					
GDP growth	0.100	0.131	0.133	0.102	0.127
Unemployment	0.096	0.110	0.099	0.108	0.128
PCE	0.096	0.111	0.104	0.108	0.119
Core PCE	0.103	0.118	0.105	0.113	0.111
Panel C: Size of the Joint Test Across Years					
GDP growth	0.108	0.127	0.193	0.106	0.122
Unemployment	0.094	0.113	0.093	0.097	0.093
PCE	0.099	0.110	0.121	0.094	0.089
Core PCE	0.105	0.113	0.107	0.111	0.085
Panel D: Size of the Joint Test Across Years and Series					
	0.061	0.236	0.190	0.191	0.226

a. The results are based on 1,000 simulations, where the bootstrap with 1,000 replications is used for the inference in each simulation.

Table 2.1: Size of the Tests

Projected Year	Bias	Autocorrelation	Wald	Signs, Positive	Signs, Consecutive
Panel A: Power of the Tests (Individual Year)					
GDP growth	0.000	0.312	0.216	0.015	0.187
Unemployment	0.000	0.326	0.236	0.008	0.207
PCE	0.000	0.375	0.255	0.012	0.221
Core PCE	0.000	0.369	0.251	0.009	0.212
Panel B: Power of the Tests (Average of Five Years)					
GDP growth	0.001	0.287	0.190	0.024	0.134
Unemployment	0.000	0.292	0.206	0.020	0.149
PCE	0.000	0.310	0.215	0.017	0.151
Core PCE	0.000	0.315	0.216	0.018	0.153
Panel C: Power of the Joint Test Across Years					
GDP growth	0.002	0.816	0.274	0.013	0.664
Unemployment	0.004	0.848	0.324	0.015	0.693
PCE	0.000	0.867	0.342	0.017	0.714
Core PCE	0.000	0.883	0.334	0.009	0.732
Panel D: Power of the Joint Test Across Years and Series					
	0.000	1.000	0.578	0.011	0.998

a. Same as Table 2.1.

Table 2.2: Power of the Tests (Independent Noise)

Projected Year	Bias	Autocorrelation	Wald	Signs, Positive	Signs, Consecutive
Panel A: Power of the Tests (Individual Year)					
GDP growth	0.003	0.283	0.210	0.008	0.179
Unemployment	0.000	0.326	0.215	0.009	0.190
PCE	0.000	0.338	0.239	0.005	0.194
Core PCE	0.001	0.349	0.262	0.010	0.192
Panel B: Power of the Tests (Average of Five Years)					
GDP growth	0.001	0.276	0.196	0.021	0.147
Unemployment	0.000	0.282	0.191	0.020	0.137
PCE	0.000	0.296	0.203	0.018	0.140
Core PCE	0.000	0.315	0.221	0.018	0.147
Panel C: Power of the Joint Test Across Years					
GDP growth	0.000	0.573	0.255	0.001	0.469
Unemployment	0.000	0.536	0.254	0.010	0.372
PCE	0.000	0.592	0.271	0.005	0.428
Core PCE	0.000	0.574	0.299	0.006	0.412
Panel D: Power of the Joint Test Across Years and Series					
	0.000	0.986	0.466	0.007	0.930

a. Same as Table 2.1.

Table 2.3: Power of the Tests (Persistent Noise)

Projected Year	Bias	Autocorrelation	Wald	Signs, Positive	Signs, Consecutive
Panel A: Power of the Tests (Individual Year)					
GDP growth	0.219	0.298	0.274	0.160	0.426
Unemployment	0.244	0.532	0.363	0.192	0.489
PCE	0.228	0.393	0.312	0.198	0.479
Core PCE	0.249	0.407	0.313	0.192	0.484
Panel B: Power of the Tests (Average of Five Years)					
GDP growth	0.222	0.264	0.284	0.191	0.363
Unemployment	0.238	0.469	0.321	0.199	0.409
PCE	0.231	0.339	0.284	0.203	0.406
Core PCE	0.246	0.363	0.290	0.209	0.409
Panel C: Power of the Joint Test Across Years					
GDP growth	0.206	0.690	0.471	0.176	0.631
Unemployment	0.234	0.825	0.475	0.197	0.499
PCE	0.242	0.817	0.483	0.219	0.575
Core PCE	0.241	0.798	0.464	0.224	0.512
Panel D: Power of the Joint Test Across Years and Series					
	0.188	0.997	0.707	0.315	0.958

a. Same as Table 2.1.

Table 2.4: Power of the Tests (Sluggish Adjustment)

Projected Year	Bias	Autocorrelation	Wald	Signs, Positive	Signs, Consecutive
Panel A: Real GDP Growth					
2009	-0.344*	0.370*	2.643	0.371	0.643
2010	-0.008	0.107	0.450	0.628	0.545
2011	-0.129	-0.058	4.432	0.335	0.404
2012	-0.185	0.146	6.054	0.236	0.627
2013	-0.138	-0.276	7.910*	0.231*	0.667
2014	-0.114	-0.185	4.669	0.322	0.392
Panel B: Unemployment Rate					
2009	0.644***	0.206	8.638*	1.000***	1.000***
2010	0.400**	0.527***	10.953**	0.833***	0.636
2011	0.250*	0.478**	8.454**	0.583	0.636
2012	0.054	0.036	0.486	0.464	0.500
2013	-0.008	0.073	0.061	0.309	0.667
2014	-0.109	-0.461	7.157	0.182*	0.600
Panel C: PCE Inflation					
2009	-0.088	0.235	0.658	0.554	0.490
2010	-0.038	0.046	0.653	0.539	0.455
2011	0.104	-0.129	2.377	0.583	0.364
2012	0.008	-0.632**	7.724*	0.494	0.320
2013	-0.050	-0.134	1.332	0.462	0.376
2014	-0.014	0.028	0.787	0.502	0.496
Panel D: Core PCE Inflation					
2009	-0.044	0.115	0.255	0.696	0.653
2010	-0.058	0.093	1.089	0.511	0.478
2011	0.029	-0.031	2.001	0.626	0.456
2012	0.023	-0.063	0.076	0.462	0.417
2013	-0.031	0.370*	3.206	0.460	0.498
2014	-0.018	-0.319	1.894	0.508	0.400
Panel E: Joint Tests Across Years					
GDP Growth	-0.153**	0.017	4.360	0.354**	0.546
Unemployment	0.205**	0.143*	5.958**	0.562	0.673**
PCE	-0.013	-0.098	2.255	0.522	0.417
Core PCE	-0.016	0.027	1.420	0.544	0.484
Panel F: Joint Test Across Years and Series					
	0.006	0.023*	3.498	0.495	0.530

a. Superscripts $*, **, ***$ denote the significance at the level of 10%, 5% and 1%, respectively. The bootstrap inference is based on 10,000 replications.

Table 2.5: Efficiency Tests of FOMC's Economic Projections (Central Tendency)

Projected Year	Bias	Autocorrelation	Wald	Signs, Positive	Signs, Consecutive
Panel A: Real GDP Growth					
2009	-0.331*	0.370*	2.469	0.500	0.429
2010	-0.004	0.008	0.064	0.616	0.640
2011	-0.150	0.078	2.964	0.376	0.453
2012	-0.162	-0.123	7.434*	0.268	0.542
2013	-0.131	-0.124	8.614*	0.272	0.582
2014	-0.132	0.113	69.335∗	0.235	0.540
Panel B: Unemployment Rate					
2009	0.656***	0.117	7.629*	0.750*	0.571
2010	0.400***	0.467**	10.763**	0.750**	0.818**
2011	0.242*	0.364*	6.692*	0.667	0.455
2012	0.077	0.139	1.406	0.580	0.545
2013	0.012	0.057	0.053	0.346	0.833**
2014	-0.114	0.186	13.909***	0.126**	0.766
Panel C: PCE Inflation					
2009	-0.063	0.262	0.583	0.443	0.571
2010	-0.037	0.210	1.222	0.493	0.635
2011	0.133	-0.213	2.339	0.464	0.507
2012	0.035	-0.185	0.568	0.732*	0.459
2013	-0.012	0.136	0.318	0.466	0.617
2014	-0.023	-0.017	0.171	0.365	0.602
Panel D: Core PCE Inflation					
2009	-0.037	0.110	0.378	0.549	0.397
2010	-0.050	-0.065	0.644	0.451	0.438
2011	0.046	-0.252	3.189	0.510	0.465
2012	0.035	0.071	0.431	0.467	0.543
2013	-0.008	0.080	0.122	0.465	0.410
2014	-0.018	-0.481	3.086	0.417	0.300
Panel E: Joint Test Across Years					
GDP Growth	-0.152**	0.054	5.147*	0.378	0.531
Unemployment	0.212**	0.222**	6.742**	0.537	0.665**
PCE	0.006	0.032	0.867	0.494	0.565
Core PCE	-0.005	-0.089	1.308	0.476	0.426
Panel F: Joint Test Across Years and Series					
	0.015	0.055**	3.516	0.471	0.547

a. Same as Table 2.5.

Table 2.6: Efficiency Tests of FOMC's Economic Projections (Range)

Projected Year	Bias	Autocorrelation	Wald	Signs, Positive	Signs, Consecutive
Panel A: Real GDP Growth					
2009	-0.723***	0.456**	8.728*	0.286	0.833**
2010	0.085	0.186	1.301	0.571	0.667
2011	-0.104	-0.134	1.148	0.300	0.444
2012	-0.071	-0.086	0.974	0.286	0.692*
2013	-0.095	-0.328	9.012**	0.200**	0.714*
2014	-0.109	-0.022	2.913	0.286	0.462
2015	-0.005	-0.813***	23.645***	0.500	0.111**
Panel B: Unemployment Rate					
2009	0.587***	0.352*	10.617*	1.000***	1.000***
2010	0.143	0.005	0.085	0.571	0.500
2011	0.045	-0.032	0.127	0.600	0.556
2012	0.043	0.002	0.071	0.571	0.462
2013	0.008	-0.003	0.050	0.333	0.571
2014	-0.051	0.113	0.564	0.214*	0.692*
2015	-0.091	0.260	3.807	0.200	0.556
Panel C: PCE Inflation					
2008	0.124	-0.228	1.120	0.714	0.500
2009	-0.089	-0.008	0.518	0.636	0.600
2010	-0.104	-0.286	5.864	0.364	0.500
2011	0.085	-0.035	0.811	0.545	0.500
2012	-0.024	-0.494	4.043	0.545	0.200
2013	-0.064	0.093	8.167	0.273	0.700
2014	-0.043	-0.052	1.629	0.200	0.778**
Panel D: Core PCE Inflation					
2008	0.039	0.041	1.492	0.714	0.667
2009	-0.053	0.052	0.680	0.455	0.500
2010	-0.093*	-0.193	7.172*	0.182*	0.700
2011	0.021	0.230	0.640	0.545	0.600
2012	-0.019	-0.038	0.185	0.364	0.400
2013	-0.045	-0.036	4.420	0.455	0.700
2014	-0.038	-0.020	2.557	0.200	0.667
Panel E: Joint Test Across Years					
GDP Growth	-0.146**	-0.106	6.817**	0.347**	0.561
Unemployment	0.098*	0.100	2.189	0.499	0.619
PCE	-0.016	-0.144	3.165	0.468	0.540
Core PCE	-0.027	0.005	2.449	0.416	0.605

a. Same as Table 2.5.

Table 2.7: Efficiency Tests of the SPF Forecast (Mean)

Projected Year	Bias	Autocorrelation	Wald	Signs, Positive	Signs, Consecutive
		Panel A: Real GDP Growth			
2009	-0.748***	0.478**	9.166*	0.286	0.833**
2010	0.073	0.145	0.884	0.571	0.667
2011	-0.075	-0.099	0.753	0.500	0.333
2012	-0.057	-0.434	4.381	0.429	0.462
2013	-0.096	-0.317	3.549	0.400	0.357
2014	-0.116	-0.237	2.684	0.357	0.385
2015	-0.000	-0.273	1.126	0.400	0.333
		Panel B: Unemployment Rate			
2009	0.597***	0.385*	11.784*	1.000***	1.000***
2010	0.121	-0.049	0.029	0.714	0.667
2011	0.034	-0.052	0.045	0.600	0.556
2012	0.029	-0.015	0.009	0.500	0.462
2013	0.009	-0.025	0.112	0.333	0.571
2014	-0.072	0.059	0.661	0.286	0.692
2015	-0.099	-0.164	3.003	0.200	0.556
		Panel C: PCE Inflation			
2008	0.109	-0.192	0.872	0.714	0.500
2009	-0.087	-0.026	0.379	0.455	0.500
2010	-0.094	-0.342	5.663	0.273	0.500
2011	0.078	-0.012	0.549	0.455	0.300
2012	-0.019	-0.288	1.063	0.545	0.400
2013	-0.073	0.543**	16.636**	0.273	0.500
2014	-0.027	-0.193	1.283	0.400	0.444
		Panel D: Core PCE Inflation			
2008	0.036	0.382*	1.719	0.714	0.667
2009	-0.064	0.111	1.008	0.364	0.500
2010	-0.087**	0.096	8.339**	0.091**	0.700
2011	0.009	0.190	0.503	0.545	0.500
2012	-0.013	0.164	0.410	0.273	0.600
2013	-0.045	0.402*	5.028	0.455	0.700
2014	-0.030	-0.026	1.198	0.100***	0.500
		Panel E: Joint Test Across Years			
GDP Growth	-0.146**	-0.105	3.220	0.420	0.481
Unemployment	0.089	0.067	2.235	0.519	0.643*
PCE	-0.016	-0.073	3.778	0.445	0.449
Core PCE	-0.028	0.189***	2.601	0.363	0.603***

a. Same as Table 2.5.

Table 2.8: Efficiency Tests of the SPF Forecast (Median)

	Nowcast	1-Year Ahead	2-Year Ahead	Cumulative
Panel A: Central Tendency				
Slope	-0.405	-0.812	0.288	-0.905
R^2	0.367	0.491	0.029	0.415
90% Conf. Int.	[-0.599, -0.166]	[-0.934, -0.277]	[-1.081, 0.745]	[-1.164, -0.421]
Panel B: Range				
Slope	-0.414	-0.720	0.112	-0.890
R^2	0.366	0.441	0.002	0.465
90% Conf. Int.	[-0.580, -0.170]	[-0.813, -0.324]	[-1.019, 1.200]	[-1.120, -0.500]

a. Analysis of cumulative revision is based on the sum of forecast revisions at all horizons.
b. Confidence interval is computed based on the block bootstrap.

Table 2.9: Regression Analysis of the FOMC's Unemployment Revisions on the GDP Growth Revisions

Chapter 3

Using Forecast Evaluation to Improve the Accuracy of the Greenbook Forecast

This chapter is a reprint from the article, "Using forecast evaluation to improve the accuracy of the Greenbook forecast," by Natsuki Arai, in *International Journal of Forecasting*, 30(1), 12-19, with permission from Elsevier.

3.1 Introduction

Much research has been concerned with forecast efficiency regressions. Recently, researchers have found evidence against efficiency in forecast efficiency regressions in a multi-horizon system. This paper uses these tests to adjust the original forecasts, and finds modest improvements for the Fed's Greenbook forecast for GDP deflator and CPI, but not for other variables.

In this paper, I propose a new method to improve the accuracy of forecasts in real time, using the results of the forecast efficiency test. Based on the evidence against the efficiency of the Greenbook forecast presented by Patton and Timmermann (2012), this paper uses the new method to adjust systematic errors of the Greenbook forecast in real time, building on the suggestion of Croushore (2012).

I find modest, but statistically significant, improvements in the out-of-sample forecast accuracy of the Greenbook forecast for GDP deflator and CPI. Since Romer and Romer (2000) showed that the Greenbook forecast is more efficient than private forecasts, the better performance of the Greenbook forecast in the 1980s has been documented in the literature. Sims (2002) confirms their result that the Greenbook forecast is more accurate than private or naïve forecasts. Faust and Wright (2009) show that the Greenbook forecast outperforms the forecasts using reduced-form models, even after giving the Greenbook forecast to these models for several quarters. In more recent periods, the Greenbook forecast does not have this advantage (Reifschneider and Tulip (2007) and Edge and Gürkaynak (2010)), but it is still an interesting benchmark for three reasons. First, the Greenbook forecast is well known in the literature and researchers have already analyzed its accuracy and efficiency from different perspectives. Second, the Greenbook forecast is the most substantial and detailed judgmental US macroeconomic forecast based on an immense range of information. Finally, as indicated by FOMC minutes and transcripts, the Greenbook forecast has played an essential role in US monetary policy.[1]

By comparing out-of-sample forecast errors during the Great Moderation, this paper shows that the adjustment based on the forecast efficiency test gives modest improvements for GDP deflator and CPI, but not for other variables. The magnitude of improvements can be as high as 18 percent in root mean

[1] For details, see the NBER working paper version of Faust and Wright (2009).

square prediction error.

The results that significant improvements can be found in the forecasts for inflation, but not in the forecasts for output growth, are consistent with a finding in the recent literature; namely, forecasts for output growth are hard to improve given a good estimate of the current state of the economy, and output growth is especially unpredictable during the Great Moderation.

The essential way in which the proposed method works is to determine if Greenbook forecasters have over- or under-reacted to incoming news in the past and then to propose a systematic adjustment for their past mistakes. Of course, judgmental forecasters are also monitoring their own performance and making these adjustments. It is possible that my correction, applied to the future Greenbook forecasts, would over-correct and effectively make the adjustment twice. However, the evidence that I show in this paper indicates that real-time implementation of my proposed adjustment would have given better out-of-sample forecasts than the Greenbook itself. That is, of course, no guarantee of future performance, but indicates that the proposed adjustment might well help.

Two methods of inference for nested forecasts—the bootstrap and the test proposed by Clark and West (2007) (henceforth referred to as the CW test), show that these improvements are statistically significant in some cases. The comparison between these methods shows that they lead to the similar results, but the CW test sometimes rejects the null even when the null model is estimated to be more accurate. This is because the CW test adjusts for the parameter estimation error. It may seem surprising that the CW test can recommend using the less accurate forecasts, but this point was made by Clark and West (2007).

Lastly, I provide extensions to the main results. First, I apply the same adjustment scheme to another subjective forecast, the SPF forecast. Second, I use a different sample period for the application to the Greenbook forecast, starting before the Great Moderation. The evidence from these two extensions is mixed, but I still find statistically significant improvements in some cases. Third, I provide an analysis in subperiods to shed some light on how the proposed adjustment improves the forecast accuracy of the Greenbook forecast.

The remainder of the paper is organized as follows: Section 3.2 describes the forecasts and vintage data I use, and Section 3.3 explains the methodology including the adjustment of forecast and inference. Section 3.4 contains the main results and possible interpretation, and Section 3.5 provides extensions of the main results. Section 3.6 concludes.

3.2 Data

This paper primarily focuses on the Greenbook forecasts for inflation and output growth: GDP deflator, CPI inflation, Core CPI inflation and GDP growth. The data for the Greenbook forecast are obtained from the Philadelphia Fed's website. Since the Greenbook forecast is the forecast prepared before the FOMC meeting, which is usually held eight times a year, I pick the forecast that is closest to the middle of each quarter to construct quarterly forecasts. Though the Greenbook forecast has been released from 1964, its forecast horizons are different, especially in early periods. With four-quarter forecast horizons, the GDP deflator and the GDP growth forecasts are available from the second quarter of 1974, the CPI forecast is available from the fourth quarter of 1979, and the Core CPI forecast is available from the first quarter of 1986.

In order to have comparability between different series and to focus on the forecasts made during the Great Moderation, I use the data from the first quarter of 1984 to the fourth quarter of 2005 as a benchmark. Given Tulip (2009)'s observation that the forecast errors made by the Greenbook forecast were largest during early in the sample period, Faust and Wright (2009) set the sample beginning in 1984 as their baseline case. I also follow this convention to prevent the volatility of the data before the Great Moderation affecting the whole analysis. Since the Greenbook forecast becomes available to the public after a lag of five years, the fourth quarter of 2005 is the most recent available data.

All forecasts and variables are quarterly, and all vintages are recorded quarterly. The vintage data are obtained from the Philadelphia Fed's website. Inflation and output growth rates are computed as annualized percent changes, $100*((\frac{x_t}{x_{t-1}})^4-1)$, where x_t is a price or output level at time t. The results using the continuously compounding annual rate of change, $400*log(\frac{x_t}{x_{t-1}})$, are listed in the online appendix, but the difference between these results is very small. For CPI and Core CPI, I use the data recorded in December 2010 and treat the data up until time $t-1$ as a vintage of time t, ignoring the issues associated with the lack of real-time data. This is because the availability of vintage data for CPI and Core CPI is limited and the revisions to these two measures are trivial. Quarterly price levels are computed by averaging the monthly values for the three months in the quarter.

Using real-time data, it is important to adopt a definition of "the realized value". In this paper, I follow Faust and Wright (2009) and treat the data released two quarters after the forecasted date as the realized value.[2] For example, I treat the output growth of the second quarter of 1994 recorded in the fourth quarter of 1994 as the realized value for the computation of forecast errors.

3.3 Method

3.3.1 Multi-Horizon Forecast Efficiency Evaluation

First, I apply the forecast efficiency evaluation across multiple horizons proposed by Patton and Timmermann (2012) to the Greenbook forecast. Let y_{t+1} be a variable at time $t+1$ to be forecasted, and $\hat{y}_{t+1|t}$ be a forecast of y_{t+1} at time t. The standard Mincer-Zarnowitz regression to test forecast efficiency is given by the following equation:

$$y_{t+1} = \alpha + \beta \hat{y}_{t+1|t} + \varepsilon_{t+1}. \tag{3.1}$$

Since standard forecast efficiency implies that forecasts are the conditional mean of forecasted variables, the null hypothesis is $[\alpha, \beta] = [0, 1]$.

Now define a revision of the forecast for t between $t-i$ and $t-j$, for $0 < i < j$, as $d_{t|i,j} \equiv \hat{y}_{t|t-i} - \hat{y}_{t|t-j}$. By definition, a recent forecast is described as the sum of the forecast at a longer horizon and subsequent forecast revisions; $\hat{y}_{t|t-i} = \hat{y}_{t|t-j} + \sum_{k=i}^{j-1} d_{t|k,k+1}$. By replacing the nowcast in Equation (3.1), $\hat{y}_{t+1|t}$, with the sum of an old forecast and subsequent revisions, $\hat{y}_{t+1|t-j} + \sum_{k=1}^{j} d_{t+1|k,k+1}$, Equation (3.1) can be rewritten as follows:

$$y_{t+1} = \alpha + \beta\ \hat{y}_{t+1|t-j} + \sum_{k=1}^{j} \gamma_k\ d_{t+1|k,k+1} + \varepsilon_{t+1}, \tag{3.2}$$

[2]For more detailed discussion, see Tulip (2009) and Faust and Wright (2009). On how data revisions affect the qualitative implication of forecasting, see Croushore (2006).

with the null hypothesis of $[\alpha, \beta, \gamma_1, \ldots, \gamma_j] = [0, 1, 1, \ldots, 1]$, where j denotes the number of forecast revisions included in the regression. This regression tests the implication of forecast efficiency that forecasts are the conditional mean and the subsequent revisions are orthogonal to the past forecasts. The F-test is used for this regression to test the null hypothesis jointly. By a Monte Carlo simulation, Patton and Timmermann (2012) show that this multi-horizon forecast efficiency evaluation has higher power to detect forecast inefficiency in finite samples.

In addition, Patton and Timmermann (2012) apply this method to the Greenbook forecast and reject its efficiency. Though Clements, Joutz, and Stekler (2007) also reject the efficiency of the Greenbook forecast by pooling forecast errors at different horizons, this multi-horizon approach is more straightforward. This paper also primarily focuses on the Greenbook forecast in the following sections.

3.3.2 Adjustment of Forecast

Given the forecast evaluation described in the previous section, I propose a simple method which is able to improve the accuracy of forecasts. Essentially, I treat the predicted value from the forecast efficiency regression as a new forecast, which adjusts the systematic errors of the original forecast.

Suppose that I evaluate forecast efficiency using the regression in Equation (3.2) every period. Specifically, I observe the following series in period t: the vintage of the forecasted variable in period t, $\{y_{j+1|t}, \ldots, y_{t|t}\}$, the Greenbook forecast $j+1$ periods before, $\{\hat{y}_{j+1|0}, \ldots, \hat{y}_{t|t-j-1}\}$, and subsequent revisions of the Greenbook forecast from $j+1$ periods before until 1 period before,

$\big\{\{d_{j+1|k,k+1}\}_{k=1}^{j}, \ldots, \{d_{t|k,k+1}\}_{k=1}^{j}\big\}$. After evaluating forecast efficiency by using these series up until time t, I plug the forecast for period $t+1$ made $j+1$ periods before and subsequent revisions, $\hat{y}_{t+1|t-j}$ and $\{d_{t+1|k,k+1}\}_{k=1}^{j}$, into the estimated equation, and treat its prediction as another forecast. In order to simplify the algorithm and to use all available information, I treat the vintage at each period as the realized value. This definition is given by the following equation:

$$\tilde{y}_{t+1|t,j} \equiv \hat{\alpha}_t + \hat{\beta}_t \ \hat{y}_{t+1|t-j} + \sum_{k=1}^{j} \ \hat{\gamma}_{t,k} \ d_{t+1|k,k+1}, \tag{3.3}$$

where $\hat{\alpha}_t, \ \hat{\beta}_t, \ \{\hat{\gamma}_{t,k}\}_{k=1}^{j}$ are the estimated coefficients in Equation (3.2) in period t. By repeating this procedure every period to obtain the predictions from the forecast efficiency regression, I form the adjusted forecast in real time.

The intuition behind this adjustment is that the new forecast adjusts the systematic errors the original forecast made in the past. If the null hypothesis in Equation (3.2) is rejected, then it means that previous forecast revisions over- or under-reacted to incoming news, or were systematically too optimistic or pessimistic, leading to inefficiencies in the original forecast. As a result, correcting these systematic errors as in Equation (3.3) gives researchers an opportunity to improve the accuracy of the original forecast. By construction, the adjusted forecasts will not contain these systematic errors, and so the adjustment cannot then be applied again. In his discussion of Patton and Timmermann (2012), Croushore (2012) suggested the idea of using the test to create improved forecasts, which I am implementing in this paper.

3.3.3 Extension to Longer-Horizon Forecasts

The extension of the baseline method to longer-horizon forecasts is straightforward: extend forecast horizon to h and focus on y_{t+h} instead of y_{t+1}. Then the extended forecast evaluation becomes

$$y_{t+h} = \alpha_h + \beta_h \ \hat{y}_{t+h|t-j} + \sum_{k=h}^{h+j-1} \gamma_{h,k} \ d_{t+h|k,k+1} + \varepsilon_{t+h} \tag{3.4}$$

with the null hypothesis of $[\alpha_h, \beta_h, \gamma_{h,h}, \ldots, \gamma_{h,h+j-1}] = [0, 1, 1, \ldots, 1]$. The only difference between Equation (3.2) and Equation (3.4) is that Equation (3.4) does not contain recent forecast revisions, $\{d_{t+h|1,2}, \ldots, d_{t+h|h-1,h}\}$, since they are not available at time t. In other words, I replace recent forecast revisions by old forecast revisions to evaluate forecast efficiency in real time. Then I define the adjusted forecast in exactly the same way:

$$\tilde{y}_{t+h|t,j} \equiv \hat{\alpha}_{h|t} + \hat{\beta}_{h|t} \ \hat{y}_{t+h|t-j} + \sum_{k=h}^{h+j-1} \hat{\gamma}_{h|t,k} \ d_{t+h|k,k+1} \tag{3.5}$$

where $\hat{\alpha}_{h|t}$, $\hat{\beta}_{h|t}$, $\{\hat{\gamma}_{h|t,k}\}_{k=h}^{h+j-1}$ are the estimated coefficients in Equation (3.4) using the series up until t. Trivially, it is a generalization of the forecast in Equation (3.3). Note that I can extend this model only for a few periods since the number of horizons in the Greenbook forecast is limited, which is four quarters ahead in my dataset.

3.3.4 Test Statistic and Inference

The natural metric to compare the forecast accuracy of the Greenbook forecast and the adjusted forecast is the out-of-sample Relative Root Mean Square Prediction Error (RRMSPE). The RRMSPE is defined as the ratio of the RMSPE of the adjusted forecast for h-period ahead, with the adjustment using j forecast revisions, to the RMSPE of the Greenbook forecast for h-period ahead:

$$RRMSPE_{h|j} \equiv \sqrt{\frac{\sum_{t=1}^{T}(\tilde{y}_{t+h|t,j} - y_{t+h})^2}{\sum_{t=1}^{T}(\hat{y}_{t+h|t} - y_{t+h})^2}}, \tag{3.6}$$

where $\tilde{y}_{t+h|t,j}$ is the adjusted forecast for $t+h$ made at t using j forecast revisions, and $\hat{y}_{t+h|t}$ is the Greenbook forecast for $t+h$ made at t, and T is the number of predictions. Since the RMSPE of the Greenbook forecast is in the denominator, the RRMSPE larger than unity indicates that the Greenbook forecast outperforms the adjusted forecast. Under the null hypothesis that the Greenbook forecast is efficient, the RRMSPE has an expected value greater than unity, because the in-sample over fitting worsens the out-of-sample predictive accuracy in small samples.

Out-of-sample RRMSPEs are calculated after forty quarters from the starting point of the sample of each series. Adjusted forecasts are computed in two ways: recursive, in which the entire sample is used to estimate the model; and rolling, in which the samples of the forty most recent quarters are used.

If I set $[\hat{\alpha}_{h|t},\ \hat{\beta}_{h|t},\ \hat{\gamma}_{h|t,h},\ \ldots,\ \hat{\gamma}_{h|t,h+j-1}]=[0,1,1,\ldots,1]$ for all t in Equation (3.5) as is consistent with the null hypothesis, the adjusted forecast becomes identical to the Greenbook forecast. In other words, the adjusted forecast nests the Greenbook forecast. When forecasting models are nested, the distribution of the test statistic presented in Diebold and Mariano (1995) is not asymptotically normal. The literature on forecast evaluation shows that testing the null hypothesis of equal MSPE for nested models with normal critical values results in severe size distortion and poor power in practice.[3] Similarly, the RRMSPE

[3]For details, see West (2006) and Faust and Wright (2013).

for nested models has a nonstandard distribution and assessing its statistical significance raises a number of econometric issues. In order to avoid these issues, this paper uses two different methods of inference: bootstrap and the CW test.

3.3.4.1 Bootstrap

The bootstrap p-values are constructed by using the null hypothesis that the Greenbook forecast is efficient and therefore it is a conditional mean of realized series. By resampling from the residuals of the AR(4) model, I first make an artificial realized series. Then I treat the conditional mean of an artificial series as artificial Greenbook forecasts and construct adjusted forecasts in exactly the same way. By computing the RRMSPE of these artificial forecasts and repeating this procedure arbitrarily many times, I can form the distribution of the bootstrap RRMSPE and report p-values of the realized RRMSPE. The specific algorithm is described in detail in Appendix C.

3.3.4.2 The CW test

This paper uses the CW test as an alternative method of inference on nested forecasts. The CW test first adjusts the noise in the MSPE due to the estimation of additional parameters in an alternative forecasting model, which nests a parsimonious null forecasting model. Then it tests the hypothesis that the null forecasting model is correctly specified and the prediction errors of these two models are the same in the population, by using normal critical values.[4]

The specific procedure is succinctly summarized in Section 2 of Clark and

[4] On the other hand, Clark and McCracken (2009) and Clark and McCracken (2011) discuss inference to test the null hypothesis that two models have the equal RMSPE in *finite sample*.

West (2007). In the context of this paper, I first compute the following statistic:

$$f_{t+h|t,j} \equiv (\hat{y}_{t+h|t} - y_{t+h})^2 - [(\tilde{y}_{t+h|t,j} - y_{t+h})^2 - (\hat{y}_{t+h|t} - \tilde{y}_{t+h|t,j})^2] \qquad (3.7)$$

where $\hat{y}_{t+h|t}$ is the Greenbook forecast for $t+h$ made at t, $\tilde{y}_{t+h|t,j}$ is the adjusted forecast for $t+h$ made at t using j revisions, and y_{t+h} is the realized value at $t+h$. Then I regress $f_{t+h|t,j}$ on a constant and derive the t-statistic. If this t-statistic is larger than 1.282 or 1.645, I reject the null hypothesis at the significance level of 10% or 5%, respectively. On the evidence from the Monte Carlo exercise in finite samples, Clark and West (2007) argue that this t-statistic is well approximated by a normal distribution, even though it has a nonstandard asymptotic distribution.

3.4 Results

Recursive and rolling RRMSPEs of adjusted forecasts are reported in Table 3.1. Adjusted forecasts are computed for nowcasts through four-quarter ahead forecasts ($h = 1, \ldots, 5$), using as many revisions as possible; for example, the adjusted nowcasts are based on four subsequent revisions, and one-period ahead adjusted forecasts are based on three subsequent revisions. The results using all the possible numbers of forecast revisions are listed in the online appendix.

3.4.1 Inflation and Output Growth

The results of the forecast accuracy of adjusted forecasts for inflation are mixed. I find significant improvements for the CPI and GDP deflator forecasts both in nowcasts and forecasts at longer horizons, whereas significant improvements for the Core CPI forecasts are only found in nowcasts.

For the CPI forecast, RRMSPEs are smaller than one at almost all horizons, both in recursive and rolling regressions. In addition, most improvements are statistically significant, the magnitude of which ranges from 4.1% to 13.0%, where the significance level varies from 1% to 10%. This implies that the Greenbook forecast made systematic errors in its forecast for CPI and adjusting it gives significant gains of forecast accuracy in real time. Also, there are many cases where I find significant improvements for the GDP deflator forecast from the adjustment, especially in rolling regressions. The magnitude of improvements ranges from 4.2% to 18.0%. However, for the GDP deflator forecast, the recursive results show few significant improvements unlike the rolling results. Unlike the case of CPI and GDP deflator forecasts, significant improvements for the Core CPI forecast are found only in nowcasts.

Faust and Wright (2009) show that the Greenbook forecast is such a good forecast for inflation that it outperforms reduced-form forecasts, even after giving Greenbook's nowcasts and longer-horizon forecasts to reduced-form forecasting models for several quarters. Sims (2002) also suggests that the superiority of the Greenbook forecast arises from its advantage in the timing of information. However, the results presented here suggest that the Greenbook forecast, (which incorporates economic judgment that perhaps make it perform better than reduced-form models), still made systematic errors and so there is still room to improve upon the Greenbook forecast.

On the other hand, the performance of the adjusted forecast for output growth is quite different from that of the inflation forecasts. I find no improvement either in recursive or rolling regressions. These results are consistent with the findings in Tulip (2009) and Faust and Wright (2009) that output growth during the Great Moderation is largely unpredictable, especially at longer horizons. Tulip (2009) argues that the predictable volatility of output growth vanishes during the Great Moderation.

3.4.2 Bootstrap and the CW Test

One objective of this paper is to compare the two different methods of inference, bootstrap and the CW test, in the context of an important practical application. As is evident from the results, the bootstrap and the CW test generally lead to the similar results. However, there are some cases where the CW test rejects the null despite the fact that the RRMSPE is larger than unity. This is because the CW test adjusts for parameter estimation error. If the restricted model is true, then it should give substantially *more* accurate forecasts in small samples, because of parameter uncertainty. If the improvement is small enough, then we could still conclude that the restricted model is to be rejected. It may seem surprising that the CW test can recommend using the less accurate forecasts, but it owes to the effects of parameter estimation error, as pointed out by Clark and West (2007).[5]

A Monte Carlo simulation reported in the online appendix confirms these results. It shows that the bootstrap inference is generally more conservative than the CW test. The CW test is modestly oversized and has higher power, whereas the bootstrap inference is modestly undersized and has lower power. Which test gives higher size-adjusted power depends on the simulations.

[5] For example, see pp. 309 of Clark and West (2007).

3.5 Extensions

3.5.1 Comparison with the SPF Forecast

In order to see whether the adjusted forecast improves upon the original forecast in the case of other judgmental forecasts, I apply the same adjustment to the Survey of Professional Forecasters (SPF) forecast for GDP deflator, CPI inflation and GDP growth. The data are obtained from the Philadelphia Fed's website.[6] The recursive and rolling results are reported in Table 3.2. More detailed results are reported in the online appendix.

The results using the SPF forecast are mixed. Unlike the case of the Greenbook forecast, I find few significant improvements either for inflation or for output growth. Even though the overall accuracy of the SPF forecast is not necessarily better than the Greenbook forecast, its errors are not as systematic as the Greenbook and the adjustments using forecast evaluation can improve the original forecast in a few but not all cases.

3.5.2 Different Subsamples

Generally, the accuracy of forecasts is very sensitive to the choice of sample periods. For example, Edge and Gürkaynak (2010) show that none of the forecasts including the Greenbook forecast or forecasts using DSGE models perform better than a constant forecast by looking at forecasts from 1992 to 2006, but that contrasts with the results from earlier samples.

In order to see how the choice of sample period affects the results, I conduct the same analysis by using the entire available sample: from the second quarter

[6] Core CPI is not included since the sample period is too short.

of 1974 for GDP deflator and output growth, and the fourth quarter of 1979 for CPI inflation. The recursive and rolling results are reported in Table 3.3. Even though there are some differences, the results of the samples from 1974 and 1979 are to some extent similar to the results of the subsample during the Great Moderation. For inflation, there are significant improvements by adjustments in both nowcasts and longer-horizon forecasts in rolling regressions. On the other hand, for output growth, I see no significant improvements using either recursive or rolling regressions.

The RRMSPEs of the adjusted SPF forecasts for the entire available samples are listed in the online appendix to conserve space, but the results are generally similar to the subsample from 1984. There are fewer improvements compared to the case of the Greenbook forecast.

3.5.3 Analysis in Subperiods

To shed some light on how the proposed adjustment improves the forecast accuracy, I provide a simple analysis of forecast revisions and a breakdown of rolling RRMSPEs in three subperiods: 1994 to 1997, 1998 to 2001, and 2002 to 2005.

First, Table 3.4 shows the sample average first-order autocorrelation and the average bias of forecast revisions for each subperiod. Since forecast revisions for the fixed target period are unpredictable under forecast efficiency, both the first-order autocorrelation and the bias of forecast revisions should be zero. However, the sample average first-order autocorrelations are negative for all subperiods and series, and average biases are sometimes notably different from zero for all series. These statistics suggest inefficiency in several subperiods. For example, average autocorrelations for CPI are -0.250, -0.293 and -0.089 in the three subperiods.

Second, Table 3.5 shows the breakdown of rolling RRMSPEs of the adjusted forecast to the Greenbook forecast for the same subperiods. The improvements vary across the series and subperiods. The improvements of the adjusted forecasts for GDP deflator and CPI inflation are mainly due to the improvements in the first and the last subperiods (1994 to 1997 and 2002 to 2005). To save space, I have listed the breakdown of recursive RRMSPEs in the online appendix.

3.6 Conclusion

This paper addresses a question related to the Fed's Greenbook forecast: Given the evidence against the forecast efficiency of the Greenbook forecast, which has been provided recently by multi-horizon forecast efficiency regressions, can researchers improve its forecast accuracy in real time? I propose a new method that uses this evidence against efficiency to adjust the Greenbook forecast. Using this method in a real-time out-of-sample forecasting exercise, I find that it gives modest improvements in the forecast accuracy of the Greenbook forecast for inflation. These improvements are statistically significant in some cases.

Specifically, I construct another forecast that adjusts systematic errors of the Greenbook forecast in real time, by collecting the predictions of a multi-horizon forecast efficiency regression every period. Then I compare out-of-sample performance of the adjusted forecast and the Greenbook forecast. By focusing on the Great Moderation, I find modest improvements from the adjustment for the GDP deflator and CPI forecasts, but not for the forecasts for other variables. The magnitude of improvement can be up to 18 percent in root mean square prediction error.

Given the results in this paper, one might be tempted to take a more general approach to the Patton and Timmermann's forecast evaluation regression, in which the coefficients shrink toward one with a Bayesian algorithm. If the prior is very dogmatic, then that would impose forecast efficiency. On the other hand, a very diffuse prior would correspond to the approach in this paper. There would be some possibilities between these two extreme approaches, but I leave this generalization as a future exercise.

Series	GDP Deflator	CPI	Core CPI	GDP Growth
Panel A: Recursive				
Nowcasts	1.033	$0.959^{***}_{\ddagger}$	$0.952^{**}_{\ddagger}$	1.026
1Q Ahead	$0.958^{*}_{\ddagger}$	1.009	1.036	1.073
2Q Ahead	1.002	$0.952_{\dagger}$	1.105	1.040
3Q Ahead	1.005	$0.906^{**}_{\ddagger}$	1.050	1.031
4Q Ahead	0.997	$0.954^{*}_{\dagger}$	1.018	1.021
Panel B: Rolling with Forty-Quarter Window				
Nowcasts	$0.877^{***}_{\ddagger}$	$0.942^{***}_{\ddagger}$	$1.001^{**}_{\ddagger}$	1.045
1Q Ahead	$0.820^{***}_{\ddagger}$	$0.906^{***}_{\ddagger}$	1.080	1.081
2Q Ahead	$0.920^{**}_{\ddagger}$	$0.899^{**}_{\dagger}$	1.133	1.065
3Q Ahead	0.984	$0.870^{***}_{\ddagger}$	1.137	$1.065_{\dagger}$
4Q Ahead	0.965	$0.912^{**}_{\dagger}$	1.085	$1.060_{\ddagger}$

a. This table shows the RRMSPEs of the adjusted forecast to the Greenbook forecast after 40 quarters from the first period. The Core CPI series start from 1986Q1 and all other series start from 1984Q1. All series end in 2005Q4.

b. Superscripts, $*$, $**$ and $***$ denote the significance at the level of 10%, 5% and 1%, respectively, based on the bootstrap with 10,000 replications. For the construction of bootstrap p-values, see Section 3.3.4.1 and Appendix C.

c. Subscripts, † and ‡, denote the significance at the level of 10% and 5%, respectively, based on the CW test. Newey-West standard errors with the lag truncation of four are used.

Table 3.1: RRMSPE of the Adjusted Forecast to the Greenbook Forecast (The Sample during the Great Moderation)

Series	GDP Deflator		CPI		GDP Growth	
	Mean	Median	Mean	Median	Mean	Median
Panel A: Recursive						
Nowcasts	1.053	1.043	$0.987^{**}_{\ddagger}$	1.017	1.001^{*}	0.996^{*}
1Q Ahead	$1.059_{\dagger}$	1.031	1.005	1.010	1.042	1.031
2Q Ahead	$1.017_{\dagger}$	1.010	1.011	1.025	1.044	1.025
3Q Ahead	1.015	1.002	1.014	1.021	1.061	1.042
4Q Ahead	$0.984_{\dagger}$	1.012	1.012	1.016	1.000	1.022
Panel B: Rolling with Forty-Quarter Window						
Nowcasts	$0.983^{**}_{\ddagger}$	1.062	$1.016^{**}_{\ddagger}$	1.053	1.207	1.169
1Q Ahead	$1.047_{\ddagger}$	1.029	1.027	1.063	1.179	1.145
2Q Ahead	$0.984_{\ddagger}$	$1.017_{\ddagger}$	1.051	1.071	1.118	1.100
3Q Ahead	$1.009_{\ddagger}$	$1.044_{\dagger}$	1.058	1.073	1.124	1.128
4Q Ahead	$1.021_{\dagger}$	1.033	1.035	1.050	1.084	1.110

a. This table shows the RRMSPEs of the adjusted forecast to the SPF forecast after 40 quarters from the first period. All series start from 1984Q1, and end in 2005Q4.
b. Same as Table 3.1.
c. Same as Table 3.1.

Table 3.2: RRMSPE of the Adjusted Forecast to the SPF Forecasts (The Sample during the Great Moderation)

Series	GDP Deflator	CPI	GDP Growth
Panel A: Recursive			
Nowcasts	1.103	1.706	1.062
1Q Ahead	1.047	$1.101_{\dagger}$	$1.023_{\ddagger}$
2Q Ahead	1.040	1.102	1.053
3Q Ahead	1.055	1.048	1.018
4Q Ahead	1.067	1.002	1.031
Panel B: Rolling with Forty-Quarter Window			
Nowcasts	$1.024^{***}_{\ddagger}$	$1.063_{\dagger}$	1.091
1Q Ahead	$0.945^{***}_{\ddagger}$	$0.973^{**}_{\ddagger}$	$1.037_{\ddagger}$
2Q Ahead	$0.993^{*}_{\dagger}$	$0.951_{\dagger}$	$1.058_{\ddagger}$
3Q Ahead	1.036	$0.856^{***}_{\ddagger}$	$1.069_{\dagger}$
4Q Ahead	1.054	$0.931^{**}_{\dagger}$	$1.076_{\dagger}$

a. This table shows the RRMSPEs of the adjusted forecast to the Greenbook forecast after 40 quarters from the first period. The GDP deflator and growth series start from 1974Q2 and the CPI series start from 1979Q4. All series end in 2005Q4.
b. Same as Table 3.1.
c. Same as Table 3.1.

Table 3.3: RRMSPE of the Adjusted Forecast to the Greenbook Forecast (The Entire Available Sample)

Series	GDP Deflator	CPI	Core CPI	GDP Growth
Panel A: Average Autocorrelation				
1994Q1-1997Q4	-0.210	-0.250	-0.206	-0.171
1998Q1-2001Q4	-0.271	-0.293	-0.205	-0.072
2002Q1-2005Q4	-0.110	-0.089	-0.134	-0.260
Panel B: Average Bias				
1994Q1-1997Q4	0.011	-0.019	-0.027	0.086
1998Q1-2001Q4	-0.039	-0.022	-0.055	-0.077
2002Q1-2005Q4	0.030	0.191	0.019	-0.177

a. This table shows the sample average first-order autocorrelation and average bias of the revisions of the Greenbook forecast, during the subperiods: 1994Q1-1997Q4 , 1998Q1-2001Q4 and 2002Q1-2005Q4.

Table 3.4: Average First-Order Autocorrelation and Average Bias of the Revisions of the Greenbook Forecast in Subperiods

Series	GDP Deflator	CPI	Core CPI	GDP Growth
Panel A: Nowcasts				
1994Q1-1997Q4	0.826	1.048	1.263	0.997
1998Q1-2001Q4	1.208	1.204	0.805	0.907
2002Q1-2005Q4	0.742	0.865	1.050	1.277
Panel B: One-Quarter-Ahead Forecasts				
1994Q1-1997Q4	0.783	1.099	1.043	1.067
1998Q1-2001Q4	1.113	1.272	1.177	1.034
2002Q1-2005Q4	0.689	0.731	1.078	1.204
Panel C: Two-Quarter-Ahead Forecasts				
1994Q1-1997Q4	0.959	1.052	1.086	1.045
1998Q1-2001Q4	1.185	1.312	0.991	0.973
2002Q1-2005Q4	0.792	0.735	1.171	1.315
Panel D: Three-Quarter-Ahead Forecasts				
1994Q1-1997Q4	1.101	0.925	1.057	1.065
1998Q1-2001Q4	1.173	1.209	0.967	0.947
2002Q1-2005Q4	0.833	0.726	1.190	1.351
Panel E: Four-Quarter-Ahead Forecasts				
1994Q1-1997Q4	1.023	0.922	0.964	1.029
1998Q1-2001Q4	1.193	1.214	1.113	0.885
2002Q1-2005Q4	0.820	0.814	1.139	1.530

a. This table shows the rolling RRMSPEs of the adjusted forecasts to the Greenbook forecast during the subperiods:1994Q1-1997Q4, 1998Q1-2001Q4 and 2002Q1-2005Q4. For Core CPI, the RRMSPEs from 1996Q1 to 1997Q4 are computed for the first subperiod.

Table 3.5: Rolling RRMSPE of the Adjusted Forecasts to the Greenbook Forecasts in Subperiods, with Forty-Quarter Window

Bibliography

ANDERSON, T. W., AND H. RUBIN (1949): "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *The Annals of Mathematical Statistics*, 20(1), 46–63.

ANG, A., G. BEKAERT, AND M. WEI (2007): "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?," *Journal of Monetary Economics*, 54(4), 1163–1212.

ARAI, N. (2014): "Using Forecast Evaluation to Improve the Accuracy of the Greenbook Forecast," *Journal International Forecasting*, 30(1), 12–19.

BALL, L., D. LEIGH, AND P. LOUNGANI (2012): "Okun's Law: Fit at 50?," Working paper.

BERNANKE, B. S. (2007): "Federal Reserve Communications," Speech, Board of Governors of the Federal Reserve.

——— (2012a): "Chairman Bernanke's Press Conference," Board of Governors of the Federal Reserve.

——— (2012b): "Recent Developments in the Labor Market," Speech, Board of Governors of the Federal Reserve.

——— (2012c): "Monetary Policy since the Onset of the Crisis," Speech.

——— (2013): "Long-Term Interest Rates," Speech.

BERNANKE, B. S., AND K. N. KUTTNER (2005): "What Explains the Stock Market's Reaction to Federal Reserve Policy?," *Journal of Finance*, 60(3), 1221–1257.

CAMPBELL, B., AND E. GHYSELS (1995): "Federal Budget Projections: A Nonparametric Assessment of Bias and Efficiency," *The Review of Economics and Statistics*, 77(1), 17–31.

——— (1997): "An Empirical Analysis of the Canadian Budget Process," *Canadian Journal of Economics*, 30(3), 553–76.

Clark, T. E., and M. W. McCracken (2009): "Nested Forecast Model Comparisons: A New Approach to Testing Equal Accuracy," Working Paper.

——— (2011): "Advances in Forecast Evaluation," Working Paper.

Clark, T. E., and K. D. West (2007): "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models," *Journal of Econometrics*, 138(1), 291–311.

Clements, M. P., F. Joutz, and H. O. Stekler (2007): "An Evaluation of the Forecasts of the Federal Reserve: A Pooled Approach," *Journal of Applied Econometrics*, 22(1), 121–136.

Coibion, O., and Y. Gorodnichenko (2010): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," NBER Working Papers 16537.

——— (2012): "What Can Survey Forecasts Tell Us about Information Rigidities?," *Journal of Political Economy*, 120(1), 116 – 159.

Croushore, D. (2006): "Forecasting with Real-Time Macroeconomic Data," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. Granger, and A. Timmermann, vol. 1, chap. 17, pp. 961–982. Elsevier.

——— (2012): "Comment on "Forecast Rationality Tests Based on Multi-Horizon Bounds"," *Journal of Business and Economic Statistics*, 30(1), 17–20.

Daly, M., and B. Hobijn (2010): "Okun's Law and the Unemployment Surprise of 2009," *FRBSF Economic Letter*, 7.

Daly, M. C., J. Fernald, Òscar Jordà, and F. Nechio (2014): "Interpreting Deviations from Okun's Law," *FRBSF Economic Letter*, 12.

D'Amico, S., and T. B. King (2013): "Flow and Stock Effects of Large-scale Treasury Purchases: Evidence on the Importance of Local Supply," *Journal of Financial Economics*, 108(2), 425 – 448.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13(3), 253–63.

Edge, R. M., and R. S. Gürkaynak (2010): "How Useful Are Estimated DSGE Model Forecasts for Central Bankers?," *Brookings Papers on Economic Activity*, vol. 2.

Eggertsson, G. B., and M. Woodford (2003): "The Zero Bound on Interest Rates and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, 34(1), 139–235.

Ellison, M., and T. J. Sargent (2012): "A Defense of the FOMC," *International Economic Review*, 53(4), 1047–1065.

Elsby, M. W. L., B. Hobijn, and A. Sahin (2010): "The Labor Market in the Great Recession," *Brookings Papers on Economic Activity*, 41(1), 1–69.

Faust, J., and J. H. Wright (2009): "Comparing Greenbook and Reduced Form Forecasts using a Large Realtime Dataset," *Journal of Business and Economic Statistics*, 27(4), 468–479.

——— (2013): "Forecasting Inflation," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. Granger, and A. Timmermann, vol. 2, chap. 3, pp. 2–56. Elsevier.

Fischer, H., M. García-Bárzana, P. Tillmann, and P. Winker (2014): "Evaluating FOMC Forecast Ranges: An Interval Data Approach," *Empirical Economics*, 47(1), 365–388.

Fujiwara, I., Y. Nakazono, and K. Ueda (2014): "Policy Regime Change against Chronic Deflation? Policy option under a long-term liquidity trap," RIETI Discussion Paper Series NO.14-E-019.

Gagnon, J., M. Raskin, J. Remache, and B. Sack (2011): "The Financial Market Effects of the Federal Reserve's Large-Scale Asset Purchases," *International Journal of Central Banking*, 7(1), 3–43.

Gavin, W. T., and G. Pande (2008): "FOMC Consensus Forecasts," *Federal Reserve Bank of St. Louis Review*, 90(3), 149–163.

Gilchrist, S., and E. Zakrajsek (2013): "The Impact of the Federal Reserve's Large-Scale Asset Purchase Programs on Corporate Credit Risk," *NBER working paper*.

Glick, R., and S. Leduc (2013): "The Effects of Unconventional and Conventional U.S. Monetary Policy on the Dollar," Working paper series, Federal Reserve Bank of San Francisco.

GÜRKAYNAK, R. S., B. SACK, AND E. SWANSON (2005): "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*, 1(1), 55–93.

HATTORI, M., AND H. S. SHIN (2008): "Yen Carry Trade and the Subprime Crisis," Working paper.

HAUSMAN, J. K., AND J. WIELAND (2014): "Abenomics: Preliminary Analysis and Outlook," *Brookings Papers on Economic Activity*, 48(2014-1).

ISIKLAR, G., K. LAHIRI, AND P. LOUNGANI (2006): "How Quickly Do Forecasters Incorporate News? Evidence from Cross-Country Surveys," *Journal of Applied Econometrics*, 21(6), 703–725.

ITO, T., AND F. S. MISHKIN (2006): "Two Decades of Japanese Monetary Policy and the Deflation Problem," in *Monetary Policy under Very Low Inflation in the Pacific Rim*, ed. by T. Ito, and A. K. Rose, vol. 15 of *NBER-EASE*, chap. 4, pp. 131–202. University of Chicago Press.

JAIN, M. (2013): "Conservatism in Inflation Forecasts," Working paper.

JOYCE, M., D. MILES, A. SCOTT, AND D. VAYANOS (2012): "Quantitative Easing and Unconventional Monetary Policy - an Introduction," *The Economic Journal*, 122(564), 271–288.

KILEY, M. T. (2013): "The Response of Equity Prices to Movements in Long-term Interest Rates Associated With Monetary Policy Statements: Before and After the Zero Lower Bound," Finance and economics discussion series, Federal Reserve.

KIMURA, T., AND J. NAKAJIMA (2013): "Identifying Conventional and Unconventional Monetary Policy Shocks: A Latent Threshold Approach," Working paper.

KLEIBERGEN, F. (2005): "Testing Parameters in GMM Without Assuming that They Are Identified," *Econometrica*, 73(4), 1103–1123.

KRISHNAMURTHY, A., AND A. VISSING-JORGENSEN (2011): "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy," *Brookings Papers on Economic Activity*, pp. 215–265.

——— (2012): "The Aggregate Demand for Treasury Debt," *Journal of Political Economy*, 120(2), 233 – 267.

——— (2013): "The Ins and Outs of LSAPs," Economic Policy Symposium, Federal Reserve Bank of Kansas City.

Kuroda, H. (2013): "Quantitative and Qualitative Monetary Easing," Speech.

Lam, W. R. (2011): "Bank of Japan's Monetary Easing Measures: Are They Powerful and Comprehensive?," *IMF Working Papers*, pp. 1–18.

Loungani, P., H. Stekler, and N. Tamirisa (2013): "Information Rigidity in Growth Forecasts: Some Cross-Country Evidence," *International Journal of Forecasting*, 29(4), 605–621.

Mishkin, F. S. (2007): "The Federal Reserve's Enhanced Communication Strategy and the Science of Monetary Policy," Speech,, Board of Governors of the Federal Reserve.

Nakazono, Y. (2013): "Strategic Behavior of Federal Open Market Committee Board Members: Evidence from Members' Forecasts," *Journal of Economic Behavior & Organization*, 93, 62–70.

Neely, C. J. (2013): "Unconventional Monetary Policy Had Large International Effects," Working paper, Federal Reserve Bank of St. Louis.

Nordhaus, W. D. (1987): "Forecasting Efficiency: Concepts and Applications," *The Review of Economics and Statistics*, 69(4), 667–74.

Patton, A. J., and A. Timmermann (2012): "Forecast Rationality Tests Based on Multi-Horizon Bounds," *Journal of Business and Economic Statistics*, 30(1).

Raskin, M. (2013a): "The Effects of Monetary Policy on Corporate Bond Yield," Working paper.

——— (2013b): "The Effects of the Federal Reserves Date-Based Forward Guidance," Finance and economics discussion series, Federal Reserve.

Reifschneider, D., and P. Tulip (2007): "Gauging the Uncertainty of the Economic Outlook from Historical Forecasting Errors," Working Paper.

Rigobon, R. (2003): "Identification through Heteroskedasticity," *The Review of Economics and Statistics*, 85(4), pp. 777–792.

Rigobon, R., and B. Sack (2003): "Measuring The Reaction of Monetary Policy to The Stock Market," *The Quarterly Journal of Economics*, 118(2), 639–669.

——— (2004): "The Impact of Monetary Policy on Asset Prices," *Journal of Monetary Economics*, 51(8), 1553–1575.

ROGERS, J. H., C. SCOTTI, AND J. H. WRIGHT (2014): "Evaluating Asset-Market Effects of Unconventional Monetary Policy: A Cross-Country Comparison," Discussion Paper.

ROMER, C. (2013): "It Takes a Regime Shift: Recent Developments in Japanese Monetary Policy through the Lens of the Great Depression," *NBER Macroeconomics Annual*, 28.

ROMER, C. D., AND D. H. ROMER (2000): "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, 90(3), 429–457.

——— (2008): "The FOMC versus the Staff: Where Can Monetary Policymakers Add Value?," *American Economic Review*, 98(2), 230–35.

ROMER, D. (2010): "A New Data Set on Monetary Policy: The Economic Forecasts of Individual Members of the FOMC," *Journal of Money, Credit and Banking*, 42(5), 951–957.

ROSA, C. (2012): "How Unconventional Are Large-Scale Asset Purchases? The Impact of Monetary Policy on Asset Prices," Staff report, Federal Reserve Bank of New York.

RÜLKE, J.-C., AND P. TILLMANN (2011): "Do FOMC Members Herd?," *Economics Letters*, 113(2), 176–179.

SHENG, X. S. (forthcoming): "Evaluating the Economic Forecasts of FOMC Members," *International Journal of Forecasting*.

SHIBAMOTO, M., AND M. TACHIBANA (2013): "The Effect of Unconventional Monetary Policy on the Macro Economy: Evidence from Japan's Quantitative Easing Policy Period," Discussion Paper Series DP2013-12, Kobe University.

SHIRATSUKA, S. (2010): "Size and Composition of the Central Bank Balance Sheet: Revisiting Japan's Experience of the Quantitative Easing Policy," *Monetary and Economic Studies*, 28, 79–106.

SIMS, C. A. (2002): "The Role of Models and Probabilities in the Monetary Policy Process," *Brookings Papers on Economic Activity*, 33(2002-2), 1–62.

STOCK, J. H., AND J. H. WRIGHT (2000): "GMM with Weak Identification," *Econometrica*, 68(5), 1055–1096.

Stock, J. H., J. H. Wright, and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20(4), 518–29.

Swanson, E. T., and J. C. Williams (forthcoming): "Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates," *The American Economic Review*.

Tillmann, P. (2011): "Strategic Forecasting on the FOMC," *European Journal of Political Economy*, 27(3), 547–553.

Tulip, P. (2009): "Has the Economy Become More Predictable? Changes in Greenbook Forecast Accuracy," *Journal of Money, Credit and Banking*, 41(6), 1217–1231.

Ueda, K. (2012a): "The Effectiveness of Non-Traditional Monetary Policy Measures: The Case of The Bank of Japan," *The Japanese Economic Review*, 63(1), 1–22.

——— (2012b): "Japan's Deflation and the Bank of Japan's Experience with Nontraditional Monetary Policy," *Journal of Money, Credit and Banking*, 44, 175–190.

Ugai, H. (2007): "Effects of the Quantitative Easing Policy: A Survey of Empirical Analyses," *Monetary and Economic Studies*, 25(1), 1–48.

West, K. D. (2006): "Forecast Evaluation," vol. 1 of *Handbook of Economic Forecasting*, chap. 3, pp. 99–134. Elsevier.

Woodford, M. (2012): "Methods of Policy Accommodation at the Interest-Rate Lower Bound," Discussion paper, Federal Reserve Bank of Kansas City.

Wright, J. H. (2012): "What does Monetary Policy do to Long-term Interest Rates at the Zero Lower Bound?," *The Economic Journal*, 122(564), 447–466.

Yellen, J. L. (2010): "The Outlook for the Economy and Monetary Policy," Speech, Federal Reserve Bank of San Francisco.

Zurawski, A., and P. D'Arcy (2009): "Japanese Retail Investors and the Carry Trade," Rba bulletin, Reserve Bank of Australia.

Appendix A

Appendix for Chapter 1

A.1 Derivation of the OLS Estimate

Describe the system of equations (1.1) and (1.2) in matrix form:

$$\begin{pmatrix} 1 & -\beta \\ -\alpha & 1 \end{pmatrix} \begin{pmatrix} \Delta i_t \\ \Delta s_t \end{pmatrix} = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} X_t + \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix}. \tag{A.1}$$

By solving this equation, we obtain the reduced-form solution of the system,

$$\begin{pmatrix} \Delta i_t \\ \Delta s_t \end{pmatrix} = \frac{1}{1-\alpha\beta} \left[\begin{pmatrix} \beta\delta + \gamma \\ \alpha\gamma + \delta \end{pmatrix} X_t + \begin{pmatrix} 1 \\ \alpha \end{pmatrix} \varepsilon_t + \begin{pmatrix} \beta \\ 1 \end{pmatrix} \eta_t \right]. \tag{A.2}$$

Let σ_X^2, σ_ε^2, and σ_η^2 as the variance of each shock. Then, the OLS estimate of α in equation (1.2) is

$$\hat{\alpha}_{OLS} = \frac{Cov(\Delta i_t, \Delta s_t)}{Var(\Delta i_t)}, \tag{A.3}$$

$$= \frac{(\beta\delta + \gamma)(\alpha\gamma + \delta)\sigma_X^2 + \alpha\sigma_\varepsilon^2 + \beta\sigma_\eta^2}{(\beta\delta + \gamma)^2\sigma_X^2 + \sigma_\varepsilon^2 + \beta^2\sigma_\eta^2}. \tag{A.4}$$

Accordingly, the bias of the OLS estimate is

$$\hat{\alpha}_{OLS} - \alpha = \frac{(1-\alpha\beta)[\delta(\beta\delta + \gamma)\sigma_X^2 + \beta\sigma_\eta^2]}{(\beta\delta + \gamma)^2\sigma_X^2 + \sigma_\varepsilon^2 + \beta^2\sigma_\eta^2}. \tag{A.5}$$

Equation (A.5) indicates that the OLS estimate has a non-zero bias, unless there is a certain restriction on parameters (such as $\alpha\beta = 1$). For the discussion about the signs of the bias in the OLS estimates, see Gilchrist and Zakrajsek (2013).

A.2 Derivation of Conditional Variance

Denote the conditional variances of the shocks on the announcement days as $\sigma^2_{X|A}$, $\sigma^2_{\varepsilon|A}$, and $\sigma^2_{\eta|A}$. Similarly, denote the conditional variances of the shocks on the non-announcement days as $\sigma^2_{X|\bar{A}}$, $\sigma^2_{\varepsilon|\bar{A}}$, and $\sigma^2_{\eta|\bar{A}}$. Given the reduced-from solution of the system in equation (A.2), the conditional variance-covariance matrices of the system, $\mathbf{\Omega}_A$ and $\mathbf{\Omega}_{\bar{A}}$, are computed as follows:

$$\mathbf{\Omega}_A = \frac{1}{(1-\alpha\beta)^2}\begin{pmatrix} (\beta\delta+\gamma)^2\sigma^2_{X|A} + \sigma^2_{\varepsilon|A} + \beta^2\sigma^2_{\eta|A} & (\beta\delta+\gamma)(\alpha\gamma+\delta)\sigma^2_{X|A} + \alpha\sigma^2_{\varepsilon|A} + \beta\sigma^2_{\eta|A} \\ \cdot & (\alpha\gamma+\delta)^2\sigma^2_{X|A} + \alpha^2\sigma^2_{\varepsilon|A} + \sigma^2_{\eta|A} \end{pmatrix}, \tag{A.6}$$

$$\mathbf{\Omega}_{\bar{A}} = \frac{1}{(1-\alpha\beta)^2}\begin{pmatrix} (\beta\delta+\gamma)^2\sigma^2_{X|\bar{A}} + \sigma^2_{\varepsilon|\bar{A}} + \beta^2\sigma^2_{\eta|\bar{A}} & (\beta\delta+\gamma)(\alpha\gamma+\delta)\sigma^2_{X|\bar{A}} + \alpha\sigma^2_{\varepsilon|\bar{A}} + \beta\sigma^2_{\eta|\bar{A}} \\ \cdot & (\alpha\gamma+\delta)^2\sigma^2_{X|\bar{A}} + \alpha^2\sigma^2_{\varepsilon|\bar{A}} + \sigma^2_{\eta|\bar{A}} \end{pmatrix}. \tag{A.7}$$

Assume that the variance of the monetary policy shock is larger on the announcement days than on the non-announcement days, but the variance of the other shocks are the same across these two sets of days. Namely, we assume that

$$\sigma^2_{\varepsilon|A} > \sigma^2_{\varepsilon|\bar{A}}, \tag{A.8}$$

$$\sigma^2_{X|A} = \sigma^2_{X|\bar{A}}, \tag{A.9}$$

$$\sigma^2_{\eta|A} = \sigma^2_{\eta|\bar{A}}. \tag{A.10}$$

When taking the difference between $\mathbf{\Omega}_A$ and $\mathbf{\Omega}_{\bar{A}}$ in equations (A.6) and (A.7), only the variances of the monetary policy shock remain and the variances of other shocks cancel out. Thus we obtain

$$\mathbf{\Omega}_A - \mathbf{\Omega}_{\bar{A}} = \frac{\sigma^2_{\varepsilon|A} - \sigma^2_{\varepsilon|\bar{A}}}{(1-\alpha\beta)^2} \begin{pmatrix} 1 & \alpha \\ \alpha & \alpha^2 \end{pmatrix}. \tag{A.11}$$

A.3 Orthogonality of Instruments

To see that the two instruments, $\mathbf{z_i}$ and $\mathbf{z_s}$, are orthogonal to the residuals $\mathbf{e}$, compute $\sum_{t=1}^{T} Z_t \cdot e_t$.

$$
\begin{aligned}
\sum_{t=1}^{T} Z_t \cdot e_t &= \mathbf{Z}' \cdot \mathbf{e}, &\text{(A.12)}\\
&= \begin{pmatrix} \mathbf{z}_i' \\ \mathbf{z}_s' \end{pmatrix} (\boldsymbol{\Delta}\mathbf{s} - \alpha\boldsymbol{\Delta}\mathbf{i}), &\text{(A.13)}\\
&= \begin{bmatrix} \frac{1}{T_A}\boldsymbol{\Delta}\mathbf{i}'_A, \ -\frac{1}{T_{\bar{A}}}\boldsymbol{\Delta}\mathbf{i}'_{\bar{A}} \\ \frac{1}{T_A}\boldsymbol{\Delta}\mathbf{s}'_A, \ -\frac{1}{T_{\bar{A}}}\boldsymbol{\Delta}\mathbf{s}'_{\bar{A}} \end{bmatrix} \left[\boldsymbol{\Delta}\mathbf{s}_A - \alpha\boldsymbol{\Delta}\mathbf{i}_A, \ \boldsymbol{\Delta}\mathbf{s}_{\bar{A}} - \alpha\boldsymbol{\Delta}\mathbf{i}_{\bar{A}}\right], &\text{(A.14)}\\
&= \begin{bmatrix} \frac{1}{T_A}\boldsymbol{\Delta}\mathbf{i}'_A \cdot (\boldsymbol{\Delta}\mathbf{s}_A - \alpha\boldsymbol{\Delta}\mathbf{i}_A) - \frac{1}{T_{\bar{A}}}\boldsymbol{\Delta}\mathbf{i}'_{\bar{A}} \cdot (\boldsymbol{\Delta}\mathbf{s}_{\bar{A}} - \alpha\boldsymbol{\Delta}\mathbf{i}_{\bar{A}}) \\ \frac{1}{T_A}\boldsymbol{\Delta}\mathbf{s}'_A \cdot (\boldsymbol{\Delta}\mathbf{s}_A - \alpha\boldsymbol{\Delta}\mathbf{i}_A) - \frac{1}{T_{\bar{A}}}\boldsymbol{\Delta}\mathbf{s}'_{\bar{A}} \cdot (\boldsymbol{\Delta}\mathbf{s}_{\bar{A}} - \alpha\boldsymbol{\Delta}\mathbf{i}_{\bar{A}}) \end{bmatrix}, &\text{(A.15)}\\
&= \begin{bmatrix} (\hat{\Omega}_{A,12} - \alpha\hat{\Omega}_{A,11}) - (\hat{\Omega}_{\bar{A},12} - \alpha\hat{\Omega}_{\bar{A},11}) \\ (\hat{\Omega}_{A,22} - \alpha\hat{\Omega}_{A,21}) - (\hat{\Omega}_{\bar{A},22} - \alpha\hat{\Omega}_{\bar{A},21}) \end{bmatrix}, &\text{(A.16)}\\
&= \begin{bmatrix} \Delta\hat{\Omega}_{12} - \alpha\Delta\hat{\Omega}_{11} \\ \Delta\hat{\Omega}_{22} - \alpha\Delta\hat{\Omega}_{21} \end{bmatrix}. &\text{(A.17)}
\end{aligned}
$$

Since $\Delta\Omega = C\begin{pmatrix} 1 & \alpha \\ \alpha & \alpha^2 \end{pmatrix}$, we have

$$\sum_{t=1}^{T} Z_t \cdot e_t \longrightarrow_p 0 \tag{A.18}$$

Accordingly, we have moment condition $E[Z_t \cdot e_t] = 0$.

Appendix B

Appendix for Chapter 2: Monte Carlo Exercise

B.1 Construction of Artificial Projections

1. Estimate a reduced-form quarterly VAR(1) of four variables, GDP growth, unemployment rate, PCE inflation and Core PCE inflation with a constant:

$$v_{t+1} = Av_t + \xi_t, \tag{B.1}$$

where

$$v_t = \begin{pmatrix} 1 \\ v_t^1 \\ \vdots \\ v_t^4 \end{pmatrix}, A = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline c_1 & a_{11} & \cdots & a_{14} \\ \vdots & \vdots & & \vdots \\ c_4 & a_{41} & \cdots & a_{44} \end{array}\right), \text{ and } \xi_t = \begin{pmatrix} 0 \\ \xi_t^1 \\ \vdots \\ \xi_t^4 \end{pmatrix}.$$

I use the vintage data of 2012 from 1984 to estimate A and the variance-covariance matrix of ξ_t. Denote estimates as $\hat{A}$ and $\hat{\Xi}$.

2. Generate the artificial realized series using $\hat{A}$ and $\hat{\Xi}$, by assuming that ξ_t is jointly normal.

3. Construct the efficient quarterly projections by iterations as in Table B.1, under the null hypothesis that the forecast is the conditional mean. I assume that the realized value is not observable until the beginning of the next period. (For example, v_1 is observable at the beginning of period 2.)

	Nowcast	1Q ahead	$\cdots$	12Q ahead
Period 1	$\hat{A}v_0$	$\hat{A}^2v_0$	$\cdots$	$\hat{A}^{13}v_0$
2	$\hat{A}v_1$	$\hat{A}^2v_1$	$\cdots$	$\hat{A}^{13}v_1$
3	$\hat{A}v_2$	$\hat{A}^2v_2$	$\cdots$	$\hat{A}^{13}v_2$
$\vdots$	$\vdots$	$\vdots$		$\vdots$

Table B.1: Simulated Quarterly Projections

4. Construct the efficient yearly projections in accordance with FOMC's projections, by assuming that period 1 is the fourth quarter of a year. Specifically, I pick the projection of the unemployment rate for the fourth quarter, as in Table B.2. Similarly, I compute the projection of GDP growth and inflation for the fourth quarter as the sum of realized values and quarterly projections, as in Table B.3.[1]

	Nowcast	Year 1	Year 2	Year 3
Period 1	$\hat{A}v_0$	$\hat{A}^5v_0$	$\hat{A}^9v_0$	$\hat{A}^{13}v_0$
2	$\hat{A}^4v_1$	$\hat{A}^8v_1$	$\hat{A}^{12}v_1$	-
3	$\hat{A}^3v_2$	$\hat{A}^7v_2$	$\hat{A}^{11}v_2$	-
4	$\hat{A}^2v_3$	$\hat{A}^6v_3$	$\hat{A}^{10}v_3$	-
5	$\hat{A}v_4$	$\hat{A}^5v_4$	$\hat{A}^9v_4$	$\hat{A}^{13}v_4$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Table B.2: Simulated Yearly Projections of the Unemployment Rate

[1] These series are in continuously compounding rate of growth.

	Nowcast	Year 1	Year 2	Year 3
Period 1	$\sum_{k=-2}^{0} v_k + \hat{A} v_0$	$\sum_{k=2}^{5} \hat{A}^k v_0$	$\sum_{k=6}^{9} \hat{A}^k v_0$	$\sum_{k=10}^{13} \hat{A}^k v_0$
2	$\sum_{k=1}^{4} \hat{A}^k v_1$	$\sum_{k=5}^{8} \hat{A}^k v_1$	$\sum_{k=9}^{12} \hat{A}^k v_1$	-
3	$v_2 + \sum_{k=1}^{3} \hat{A}^k v_2$	$\sum_{k=4}^{7} \hat{A}^k v_2$	$\sum_{k=8}^{11} \hat{A}^k v_2$	-
4	$\sum_{k=2}^{3} v_k + \sum_{k=1}^{2} \hat{A}^k v_3$	$\sum_{k=3}^{6} \hat{A}^k v_3$	$\sum_{k=7}^{10} \hat{A}^k v_3$	-
5	$\sum_{k=2}^{4} v_k + \hat{A} v_4$	$\sum_{k=2}^{5} \hat{A}^k v_4$	$\sum_{k=6}^{9} \hat{A}^k v_4$	$\sum_{k=10}^{13} \hat{A}^k v_4$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Table B.3: Simulated Yearly Projections of GDP Growth and Inflation

B.2 Computation of Forecast Revisions (Size)

Based on the efficient yearly projections, I compute the revisions of the unemployment rate projections as in Table B.4, and compute the revisions of GDP growth and inflation projections as in Table B.5. Then, I apply the tests in this paper to these revisions. By repeating the whole exercise many times, I report the probability of rejections as the size of the tests.

	1st Year	2nd Year	3rd Year	$\cdots$
1st Period	$\hat{A}^4 v_1 - \hat{A}^5 v_0$	$\hat{A}^8 v_1 - \hat{A}^9 v_0$	$\hat{A}^{12} v_1 - \hat{A}^{13} v_0$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	
4th Period	$\hat{A} v_4 - \hat{A}^2 v_3$	$\hat{A}^5 v_4 - \hat{A}^6 v_3$	$\hat{A}^9 v_4 - \hat{A}^{10} v_3$	$\cdots$
$\vdots$		$\vdots$	$\vdots$	
8th Period		$\hat{A} v_8 - \hat{A}^2 v_7$	$\hat{A}^5 v_8 - \hat{A}^6 v_7$	$\cdots$
$\vdots$			$\vdots$	
12th Period			$\hat{A} v_{12} - \hat{A}^2 v_{11}$	$\cdots$

Table B.4: Simulated Revisions of the Unemployment Projections

	1st Year	2nd Year	3rd Year	$\cdots$
1st Period	$\sum_{k=1}^{4} \hat{A}^k v_1 - \sum_{k=2}^{5} \hat{A}^k v_0$	$\sum_{k=5}^{8} \hat{A}^k v_1 - \sum_{k=6}^{9} \hat{A}^k v_0$	$\sum_{k=9}^{12} \hat{A}^k v_1 - \sum_{k=10}^{13} \hat{A}^k v_0$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	
4th Period	$v_4 + \hat{A} v_4 - \sum_{k=1}^{2} \hat{A}^k v_3$	$\sum_{k=2}^{5} \hat{A}^k v_4 - \sum_{k=3}^{6} \hat{A}^k v_3$	$\sum_{k=6}^{9} \hat{A}^k v_4 - \sum_{k=7}^{10} \hat{A}^k v_3$	$\cdots$
$\vdots$		$\vdots$	$\vdots$	
8th Period		$v_8 + \hat{A} v_8 - \sum_{k=1}^{2} \hat{A}^k v_7$	$\sum_{k=2}^{5} \hat{A}^k v_8 - \sum_{k=3}^{6} \hat{A}^k v_7$	$\cdots$
$\vdots$			$\vdots$	
12th Period			$v_{12} + \hat{A} v_{12} - \sum_{k=1}^{2} \hat{A}^k v_{11}$	$\cdots$

Table B.5: Simulated Revisions of the GDP Growth and Inflation projections

B.3 Construction of Inefficient Projections (Power)

To compute the power of the tests, construct the three types of inefficient forecasts. Denote the efficient forecast constructed in Section B.1 as $\hat{y}^{b,*}_{t+h|t+j}$, for time $t+h$ made at time $t+j$ for $0 < j < h$.

1. The forecast with the independent noise is computed as follows:

$$\hat{y}^{b,I}_{t+h|t+j} = \hat{y}^{b,*}_{t+h|t+j} + \varepsilon_{t+h|t+j}, \tag{B.2}$$

   where $\varepsilon_{t+h|t+j}$ is an independent white noise.

2. The forecast with the persistent noise across multiple horizons are computed as follows:

$$\hat{y}^{b,P}_{t+h|t+j} = \hat{y}^{b,*}_{t+h|t+j} + \eta^{h-j-1}\varepsilon_{t+j}, \tag{B.3}$$

   where η is the parameter of the persistence in the noise such that $0 < \eta < 1$, and ε_{t+j} is an independent white noise. The forecaster receives an independent noise every period, but this noise affects all forecasts at different horizons. I set $\eta = 0.8$ in the simulation.

3. The forecasts with a sluggish adjustment are computed as follows:

$$\hat{y}^{b,S}_{t+h|t+j} = \delta\hat{y}^{b,*}_{t+h|t+j} + (1-\delta)\hat{y}^{b,*}_{t+h|t+j-1}, \tag{B.4}$$

   where δ is the parameter of the sluggishness in the adjustment such that $0 < \delta < 1$. The forecast is computed as the weighted average of efficient forecasts in the current period and the previous period. I set $\delta = 0.5$ in the simulation.

Given these inefficient forecasts, I apply the tests in this paper to the revisions of these forecasts. By repeating the whole exercise many times, I report the probability of rejections as the power of the tests.

Appendix C

Appendix for Chapter 3: Construction of Bootstrap P-values

The algorithm to construct bootstrap p-values is described as follows:

1. Fit AR(4) model to realized series, $\{y_t\}$;

$$y_t = \alpha + \sum_{k=1}^{4} \phi_k \; y_{t-k} + \varepsilon_t. \tag{C.1}$$

2. Randomly resample residuals and create artificial sample, $\{y_t^b\}$;

$$y_t^b = \hat{\alpha} + \sum_{k=1}^{4} \hat{\phi}_k \; y_{t-k}^b + e_t^b, \tag{C.2}$$

   where $\hat{\alpha}$ and $\{\hat{\phi}_k\}_{k=1}^4$ are estimated coefficients of AR(4) model in equation (C.1), and e_t^b is a randomly resampled residual. I randomly pick a block of four observations to set the initial observations of an artificial sample.

3. Calculate the conditional mean of an artificial sample at all horizons and take it as an artificial Greenbook forecast. Specifically, the conditional

mean is computed in the following way:

$$\hat{y}^b_{t+h|t} = \begin{cases} \hat{\alpha} + \sum_{k=1}^{4} \hat{\phi}_k \; y^b_{t-k} & \text{if } h = 1 \\ \hat{\alpha} + \sum_{k=1}^{h-1} \hat{\phi}_k \; \hat{y}^b_{t+h-k|t} + \sum_{k=h}^{4} \hat{\phi}_k \; y^b_{t+h-k} & \text{if } 2 \leq h \leq 4 \\ \hat{\alpha} + \sum_{k=1}^{4} \hat{\phi}_k \; \hat{y}^b_{t+h-k|t} & \text{if } h = 5 \end{cases}$$

For the first four observations, I take the unconditional mean as the forecasts at all horizons.

4. Given an artificial sample and artificial Greenbook forecast, construct an adjusted forecast in exactly the same way in Section 3.3 for the h-period ahead forecast:

$$\tilde{y}^b_{t+h|t,j} \equiv \hat{\alpha}^b_{h|t} + \hat{\beta}^b_{h|t} \; \hat{y}^b_{t+h|t-j} + \sum_{k=h}^{h+j-1} \; \hat{\gamma}^b_{h|t,k} \; d^b_{t+h|k,k+1} \tag{C.3}$$

where $d^b_{t|i,j} \equiv \hat{y}^b_{t|t-i} - \hat{y}^b_{t|t-j}$ for $0 < i < j$ and $\hat{\alpha}^b_{h|t}$, $\hat{\beta}^b_{h|t}$, $\{\hat{\gamma}^b_{h|t,k}\}_{k=h}^{h+j-1}$ are the coefficients from the forecast efficiency regression using artificial data up until t. I plug the forecast for period $t+h$ made $h+j$ period before and subsequent revisions, $\hat{y}^b_{t+h|t-j}$ and $\{d^b_{t+h|k,k+1}\}_{k=h}^{h+j-1}$, into the estimated equation and treat its prediction as another forecast.

5. Repeat this procedure every period to obtain predictions. I take these predictions as the artificial adjusted forecast.

6. Compute the RRMSPE of the artificial adjusted forecast to the artificial Greenbook forecast. Unlike the procedure in Section 3.3, I assume that the realized series are observable and never revised.

7. Repeat the steps 2 to 6 to form the distribution of the bootstrap RRMSPE. I report p-values of the realized RRMSPE according to this distribution.

Curriculum Vitae

Natsuki Arai received the B.A. and M.A degree in Economics from University of Tokyo, Japan in 2003 and 2005, respectively. Between 2005 and 2008, he worked as a Japanese government official in the Japanese Ministry of Fiance. He enrolled in the Ph.D. program in Economics at the Johns Hopkins University in 2008. After completing the M.A. degree in Economics, he worked at the Japanese Tax Agency and the Ministry of Finance, but restarted his research at the Johns Hopkins University in 2013. He will start his work at National Chengchi University in Taiwan from February 2015.